

### WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



### INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5: WO 93/13663 (11) International Publication Number: A01N 43/22, A61K 31/71 A1 22 July 1993 (22.07.93) C07H 17/08, C12N 1/21 (43) International Publication Date: C12P 19/62 (81) Designated States: AU, CA, JP, KR, European patent (AT, (21) International Application Number: PCT/US92/00427 BE, CH, DE, DK, ES, FR, GB, GR, IT, LU, MC, NL, 17 January 1992 (17.01.92) SE). (22) International Filing Date: **Published** (71) Applicant: ABBOTT LABORATORIES [US/US]; CHAD D-377/AP6D-2, One Abbott Park Road, Abbott Park, IL With international search report. 60064-3500 (US). (72) Inventors: KATZ, Leonard; 844 North Avenue, Waukegan, IL 60085 (US). DONADIO, Stefano; 110 Brookhill Road, Libertyville, IL 60048 (US). MCALPINE, James, B.; 211 West Rockland Road, Libertyville, IL 60048 (US). (74) Agents: GORMAN, Edward, Hoover, Jr. et al.; Abbott Laboratories, CHAD D-377/AP6D-2, One Abbott Park Road, Abbott Park, IL 60064-3500 (US).

(54) Title: METHOD OF DIRECTING BIOSYNTHESIS OF SPECIFIC POLYKETIDES

#### (57) Abstract

A method to produce novel polyketide structures by designing and introducing specified changes in the DNA governing the synthesis of the polyketide is disclosed. The biosynthesis of specific polyketide analogs is accomplished by genetic manipulation of a polyketide-producing microorganism by isolating a polyketide biosynthetic gene-containing DNA sequence, identifying enzymatic activities associated within the DNA sequence, introducing one or more specified changes into the DNA sequence which codes for one of the enzymatic activities which results in an altered DNA sequence, introducing the altered DNA sequence into the polyketide-producing microorganism to replace the original sequence, growing a culture of the altered microorganism under conditions suitable for the formation of the specific polyketide analog, and isolating the specific polyketide analog from the culture. The method is most useful when the segment of the chromosome modified is involved in an enzymatic activity associated with polyketide biosynthesis, particularly for manipulating polyketide synthase genes from Saccharapolyspora or Streptomyces.

## -

### FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Applria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi Nafaria
BB	Barbados	GB	United Kingdom	NL	Netherlands
8E	Belgium	GN	Guinea	NO	Norway
	Burkina Faso	GR	Greece	NZ	New Zealand
BF		HU	Hungary	PŁ	Poland
BG	Bulgaria	IE	Ireland	PT	Portugal
BJ	Benin	£T.	Italy	RO	Romania .
BR	Brazil	JP	Japan-	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic	SD	Sudan
CF	Central African Republic	121	of Korea	SE	Sweden
CC	Coago	1/0	Republic of Korea	SK	Slovak Republic
CH	Switzerland	KR	Kazakhstan	SN	Senegal
CI	Côte d'Ivoire	ΚZ	Liechtenstein	SU	Soviet Union
CM	Cameroon	IJ		TD	Chad
cs	Czechoslovakia -	LK	Sri Lanka	TG	Тово
CZ	Czech Republic	เม	Lincobourg	UA	Ukraine
DE	Germany	MC	Monaco	US	United States of America
DК	Denmark	MG	Madagascar	VN	Viet Nam
ES	Spain	ML	Mali	* 13	VICE I VIII
FI	Finland	MN	Mongolia		

15

20

25

30

35

### METHOD OF DIRECTING BIOSYNTHESIS OF SPECIFIC POLYKETIDES

### Field of the Invention.

The present invention relates to a method for directing the biosynthesis of specific polyketide analogs by genetic manipulation. In particuar, polyketide biosynthetic genes are manipulated to produce precise, novel polyketides of predicted structure.

#### 10 Background of the Invention

Polyketides are a large class of natural products that includes many important antibiotics and immunosuppressants such as erythromycins, tetracyclines, and rapamycins. Their synthesis proceeds by an ordered condensation of acyl esters to generate carbon chains of varying length and substitution pattern that are later converted to mature polyketides. This process has long been recognized as resembling fatty acid biosynthesis, but with important differences. Unlike a fatty acid synthase, a typical polyketide synthase is programmed to make many choices during carbon chain assembly: For example, the choice of "starter" and "extender" units, which are often selected from acetate, propionate or butyrate residues in a defined sequence. The choice of using a full cycle of reductiondehydration-reduction after some condensation steps, omitting it completely, or using one of two incomplete cycles (reduction alone or reduction followed by dehydration), which determines the pattern of keto or hydroxyl groups and the degree of saturation at different points in the chain is additionally programed. Finally the choice of stereochemistry for the substituents at many of the carbon atoms is programmed by the polyketide synthase.

Because of the commercial significance of *Streptomyces*, a great amount of effort has been expended in the study of *Streptomyces* genetics. Consequently much is known about *Streptomyces* and several cloning vectors exist for transformations of the organism.

Although many polyketides have been identified, there remains the need to obtain novel polyketide structures with enhanced properties. Current methods of obtaining such molecules include screening of natural isolates and chemical modification of existing polyketides, both of which are costly and time consuming. Current screening methods are based on gross properties of the molecule, i.e. antibacterial, antifungal

7.7 7.7

activity, etc., and both *a priori* knowledge of the structure of the molecules obtained or predetermination of enhanced properties are virtually impossible. Chemical modification of preexisting structures has been successfully employed, but it still suffers from practical limitations to the type of compounds obtainable, largely connected to the poor yield of multistep syntheses and available chemistry to effect modifications. The following modifications are extremely difficult or inefficient at the present time: change of the stereochemistry of the side chains in the completed polyketide; change of the length of the polyketide by removal or addition of carbon units from the interior of the acyl chain; and dehydroxylation at unique positions in the acyl chain. Accordingly, there exists the need to obtain molecules wherein such changes can be specified and performed and would represent an improvement in the technology to produce altered polyketide molecules with predicted structure.

15

20

25

30

35

10

### Summary of the Invention

The present invention provides a method to produce novel structures from designing and introducing specified changes in the DNA governing the synthesis of the polyketide. According to the method of the present invention, the biosynthesis of specific polyketide analogs is accomplished by genetic manipulation of a polyketide-producing microorganism comprising the steps of:

- (1) isolating a polyketide biosynthetic gene-containing DNA sequence;
  - (2) identifying enzymatic activities associated within said DNA sequence;
  - (3) introducing one or more specified changes into said DNA sequence which codes for one of said enzymatic activities which results in an altered DNA sequence;
  - (4) introducing said altered DNA sequence into the polyketideproducing microorganism to replace the original sequence;
  - (5) growing a culture of the altered microorganism under conditions suitable for the formation of the specific polyketide analog; and
    - (6) isolating said specific polyketide analog from the culture.

The present method is most useful when the segment of the chromosome modified is involved in an enzymatic activity associated with polyketide biosynthesis. The present invention is especially useful



-

in manipulating polyketide biosynthetic genes from Streptomyces, an organism which provides over one-half of the clinically useful antibiotics.

### Brief Description of the Drawings

FIG. 1 illustrates the organization of gene encoding polyketide 5 synthase and designated eryA as follows: (a) Map coordinates of the DNA; (b) DOTPLOT of the output of COMPARE (window = 50, stringency = 32) program (Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin, Biotechnology Center, 1710 University Avenue, Madison, Wisconsin 53705 of eryA segment (x-axis) vs. 10 subsegment of eryA comprises between 23 - 27.5 sequence coordinates (yaxis) [see Fig. 2]; (c) Open reading frame organization of eryA and enzymatic activities encoded. PT = propionyltransferase; ACP = acyl carrier protein;  $KS=\beta$ -ketoacyl ACP synthase; RmT=(2R) methylmalonyl CoA transferase;  $KR = \beta$ -ketoredu mase; SmT = (2S) methylmalonyl CoA 1.5 transferase; DH = dehydratase; EK = enoylreductase; TE = thioesterase; and (d) Schematic diagram showing the extent of each of the six modules in eryA.

FIG. 2. illustrates the nucleotide sequence of *eryA* with corresponding translation of the three open reading frames. Standard one letter codes for the amino acids appear beneath their respective nucleic acid codons. The standard one letter codes for the amino acid sequences are as follows:

-alanine 25 A -arginine R -asparagine N -aspartic acid D C -cysteine -glutamine 30 Q -glutamic acid E -glycine G -histidine H -isoleucine I -leucine 35 L K -lysine -methionine (start) M -phenylalanine F

PCT/US92/00427 WO 93/13663

4

-proline P S -serine -threonine -tryptophan W -tyrosine Y

-valine

V

44

5

30

FIG. 3. is a schematic representation of Type I, Type II and Type III changes in eryA and structures of corresponding novel polyketides produced.  $\Delta 69$  (Type I) and  $\Delta 33$  (Type II) represent in-frame deletions of 10 the base pairs in the DNA segments corresponding to the KR of module 2 and the  $\beta$ -ketoacyl ACP synthase of module 2, respectively. Insertion of a complete copy of module 4 within module 1 is also shown. Production of 11-epifluoro-15-norerythromycin in strain that carries  $\Delta 33$  occurs when substrate analog (2S,3S,4S,5S)2,4-dimethyl-3-fluoro-5-hydroxyhexanoic acid-ethyl thioester is fed.

FIG. 4 illustrates the restriction site coordinates of cosmid pR1 5' to the sequence of eryA (Fig 2).

20 Detailed Description of the Invention

For the purposes of the present invention as disclosed and claimed herein, the following terms are defined.

The term "polyketide" as used refers to a large and diverse class of natural products, including antibiotics, pigments, and 25 immunosuppressants. Antibiotics include, but are not limited to anthracyclines, tetracyclines, polyethers, ansamycins, macrolides of different types (polyenes and avermectins as well as classical macrolides such as erythromycins).

The term "polyketide-producing microorganism" as used herein includes any Actinomycetales which can produced a polyketide. Examples of Actinomycetes that produce polyketides include but are not limited to Micromonospora rosaria, Micromonospora megalomicea, Sacharapolyspora erythraea, Streptomyces antibioticus, Streptomyces albireticuli, Streptomyces ambofasciens, Streptomyces avermitilis, 35 Streptomyces fradiae, Streptomyces hygroscopicus, Streptomyces tsukubaensis, Streptomyces griseus, Streptomyces mycarofasciens, Streptomyces platensis, Streptomyces venezuelae, Streptomyces

10

15

20

25

35

4

violaceoniger, and various Actinomadura, Dactylosporangium and Nocardia strains that produce polyether type of polyketides.

The term "polyketide synthase" as used herein refers to the complex of enzymatic activities responsible for the biosynthesis of polyketides which include but are not limited to  $\beta$ -ketoreductase, dehydratase, acyl carrier protein, enoylreductase,  $\beta$ -ketoacyl ACP synthase, and acyltransferase.

The term "extender" as used herein refers to a coenzyme A thioester of a dicarboxylate which is incorporated into a polyketide by a polyketide synthase.

The term "starter" as used herein refers to a coenzyme A thioester of a carboxylic acid which is used by the polyketide synthase as the first building block of the polyketide.

The term "eryA" as used herein refers to the genes involved in the formation of the polyketide moiety of erythromycin.

The term "condensation" as used herein refers to the addition of an extender unit out to the nascent polyketide chain and requires the action of  $\beta$ -ketoacyl ACP synthase, acyltransferase, and acyl carrier protein.

The term " $\beta$ -carbonyl processing" as used herein refers to changes effecting the carbonyl group of the growing polyketide via  $\beta$ -ketoreductase, dehydratase, and enoylreductase.

The term "module" as used herein refers to the genetic element encoding one condensation step, as defined above, and one  $\beta$ -carbonyl processing step, as defined herein.

The term "Type I change" as used herein refers to changes in DNA sequence which will result in the production of polyketide rings of length identical to that of 6-deoxyerythronolide A, but with altered functional groups at specific ring positions.

The term "Type II change" as used herein refers to alterations
which will result in the production of macrolide rings only when fed
exogenously with substrate analogs, e.g.thioesters of appropriate acyl
compounds of various length. Thus Type II mutants are erythromycin
non-producing (Ery-) mutants. The structure of the resulting macrolides
will depend on the substrate employed.

The term "Type III change" as used herein refers to alterations which will result in the biosynthesis of macrolide rings of length reduced (deletion) or increased (insertion) by two carbon units, or macrolide rings altered in specific portions of the chain (replacement).

20

25

In its broadest sense, the present invention entails a general procedure for producing novel polyketide structures in vivo by selectively altering the genetic information of the organism that naturally produces a related polyketide. A set of examples described herein are a series of novel polyketides that make use of the genetic information for the biosynthesis of the polyketide portion of the macrolide antibiotic erythromycin. The organization of the segment of the Saccharapolyspora erythraea chromosome, designated eryA, and the corresponding polypeptides which it encodes that determine the biosynthesis of the polyketide segment of erythromycin, are shown in FIG. 1. It is seen that eryA is organized in modules, as shown, and that each module takes care of one condensation step, through the action of the  $\beta$ -ketoacyl ACP synthase specified within, wherein an extender unit, methylmalonyl CoA, is added first to the starter unit, propionyl CoA, and then to the successively growing acyl chain. The precise succession of elongation steps is dictated by the genetic order of the six modules: module 1 determines the first condensation; module 2, the second; module 3, the third, and so on until the sixth condensation step has occurred. Furthermore, the processing of the growing chain after each condensation is also determined by the information within each module. Thus  $\beta$ ketoreduction of the β-carbonyl takes place after each step except for step 3, as determined by the presence of a functional \beta-ketoreductase in all modules except module 3, whereas dehydration and enoyireduction only take place after the fourth extender unit is added to the growing acyl chain, as determined by the presence of dehydratase and enoyireductase in module 4. Furthermore, the choice of the correct enantiomer (2R or 2S) of methylmalonyl-CoA as the extender unit employed at each condensation is specified by the acyltransferase function determined by each module (FIG. 1C).

In the present invention, novel polyketide molecules of desired structure are produced by the introduction of specific genetic alterations of the eryA sequence into the Sac. erythraea chromosome. The complete nucleotide sequence of the eryA segment of the Sac. erythraea chromosome and the sequence of the corresponding polypeptides are shown in FIG. 2. Three types of alterations to the eryA DNA sequence are described: (i) those inactivating a single function in a module which does not arrest acyl chain growth (β-ketoreductase, dehydratase, or enoylreductase); (ii) those inactivating a single function in a module

10

15

3

AN M

which does arrest chain growth (\beta-ketoacyl ACP synthase, acyltransferase or acyl carrier protein); and (iii) those affecting an entire module (deletion, insertion, or replacement). The novel polyketides produced by strains carrying these types of mutations can be classified accordingly. Type I changes will result in the production of polyketide rings of length identical to that of 6-deoxyerythronolide A, but with altered functional groups at specific ring positions. Strains carrying type II alterations will result in the production of macrolide rings only when fed exogenously with substrate analogs, e.g.thioesters of appropriate acyl compounds of various length. Thus Type II mutants are erythromycin non-producing (Ery-) mutants. The structure of the resulting macrolides will depend on the substrate employed. Type III changes will result in the biosynthesis of macrolide rings of length reduced (deletion) or increased (insertion) by two carbon units, or macrolide rings altered in specific portions of the chain (replacement). A schematic representation of some examples of Type I, Type II and Type III alterations in eryA and the corresponding novel polyketides produced in hosts that carry such alterations is shown in FIG. 3.

In the examples described herein, specific mutations in the eryA region of the Sac. erythraea chromosome are introduced by a simple two-20 step approach: 1) introduction of a specified change in a cloned DNA segment; 2) exchange of the wild type allele with the mutated one. Step 1 requires standard recombinant DNA manipulations employing E. coli as the host. Step 2 requires one or more plasmids out of the several E. coli-Sac. erythraea shuttle vectors available and a simple screening procedure 25 for the presence of the colony carrying the altered gene. Two methods are used to introduce the altered allele into the chromosome to replace the wild type allele. The first employs gene replacement, described in Examples 7, 11, 15, 19 and 24, wherein the gene to be altered, along with adjacent upstream and downstream DNA, is mutated and cloned into a 30 Sac. erythraea non-replicating vector. The plasmid carrying the altered allele is then introduced into the host strain by transformation of protoplasts employing selection for a plasmid marker. Since the plasmid does not replicate, regenerated cells that carry the marker have undergone a single homologous recombination between one of the two segments 35 flanking the mutation on the plasmid and its homologous counterpart in the chromosome. Some of the colonies that have subsequently lost the marker will have undergone a second recombination between the other

F.

plasmid borne adjacent DNA segment and its homologous chromosomal counterpart resulting in the retention of the mutation in the chromosome, replacing the normal allele with the mutant one. The second method to introduce an altered allele into the chromosome employs gene conversion, described in Examples 37 and 43. In this method, an Ery Sac. erythraea strain carrying a deletion of a specified region of the eryA segment of the chromosome is used as a host. Into a Sac. erythraea multicopy plasmid that carries a selectable marker is cloned the wild type counterpart (segment 1) of the eryA segment mutant in the host. Subsequently, the desired homologous or heterologous DNA 10 segment to be introduced (segment 2) is cloned within the portion of segment 1 which is deleted in the mutant strain. The resulting plasmid is then introduced into the host employing selection for the marker. Among the transformants will be a population that have integrated segments 1 and 2 from the plasmid by the process of gene conversion 15 which can be verified by examination of the DNA among the colonies that have recovered the ability to produce erythromycin.

Two examples each of Types I, II and III alterations to the eryA DNA sequence and the resultant novel polyketides produced are described in the examples described herein. Examples 1 through 8, 9 through 12 and 13 20 through 16 describe the construction and effect of three Type I mutants. Examples 17 through 22 and 23 through 27 describe the construction of two Type II mutants and the effects of feeding two different synthetic substrates to the mutant strains. Examples 28 through 38 and 39 through 44 outline the steps in constructing Type III changes and their respective 25 effects on the structure of the novel polyketides produced. In Examples 1 through 7 a plasmid that contains a substantial deletion of the segment of the gene corresponding to the b-ketoreductase of module 5 is created, the altered gene is inserted into the Sac. erythraea chromosome to replace the wild type allele and the new strain carrying the altered gene is identified 30 and isolated. In Example 8, the new strain is fermented and the novel polyketide 5-oxo-5,6-dideoxy-3α-mycarosyl erythronolide B that results from the introduction of the mutant allele is isolated. In Examples 9 through 11, a mutation is introduced into the  $\beta$ -ketoreductase of module 2 and the mutated allele is then used to replace the wild type allele in the 35 chromosome. In Example 12, the strain carrying the altered allele is fermented and the novel compound 11-oxo-11-deoxyerythromycin A is isolated. Similarly, in Examples 13 through 16 a mutation is introduced



into the dehydratase of module 4 and the mutated allele is then used to replace the wild type allele in the chromosome. The strain carrying this altered allele is then fermented and the novel products 7hydroxyerythromycin A and 6-deoxy-7-hydroxyerythromycin A are isolated. In Examples 17 through 21, a mutation is made in the DNA corresponding to the  $\beta$ -ketoacyl-ACP synthase of module 1 and introduced into the chromosome to replace the wild type allele. This mutation has the effect of arresting the synthesis of the polyketide chain and results in the Ery phenotype. The synthetic substrate (25,3R,45,5S)3,5-dihydroxy-2,4dimethylhexanoic acid-ethyl ester is then made and fed to the mutant 10 resulting in the production of the novel compound (14S,15S)14(1hydroxyethyl)erythromycin. Similarly, in Examples 22 through 24, a mutation is created in the  $\beta$ -ketoacyl-ACP synthase of module 2 and introduced into the chromosome to replace the wild type allele. In Example 25 and 26, the synthetic substrate (25,35,45,55)2,4-dimethyl-3-15 fluoro-5-hydroxyhexanoic acid-ethyl thioester is made and fed to the module 2 β-ketoacyl-ACP synthase mutant and the resulting novel compound 11-epifluoro-15-norerythromycin is isolated. In Examples 27 through 38, a copy of the DNA sequence corresponding to module 4 is introduced into the deleted segment of the  $\beta\text{-ketoacyl-ACP}$  synthase of 20 module 1 resulting in the production of the novel compound 14(1propyl)erythromycin. In Examples 40 through 44, a copy of the DNA sequence corresponding to module 5 is introduced into the deleted segment of the β-ketoacyl ACP synthase of module 1 resulting in the production of the novel compound 14[1(1-hydroxypropyl)]erythromycin. 25

### **GENERAL METHODS**

### 30 Materials, Plasmids and Bacterial Strains

35

Restriction endonucleases, T4 DNA ligase, nick-translation kit, competent  $\underline{E}$ .  $\underline{coli}$  DH5 $\alpha$  cells , X-gal, IPTG, and plasmids pUC19 and pUC12 are purchased from Bethesda Research Laboratories (BRL), Gaithersburg, MD. [ $\alpha$ -32P]dCTP and Hybond N are from Amersham Corp., Chicago, IL. Seakem LE agarose and Seaplaque low gelling temperature agarose are from FMC Bioproducts, Rockland, ME.  $\underline{E}$ .  $\underline{coli}$  K12 strains carrying the  $\underline{E}$ .  $\underline{coli}$ -Sac. shuttle plasmids pWHM3 or pWHM4 (Vara et al.,  $\underline{I}$ . Bacteriol., 171: 5872 (1989)) or the cosmids pS1 (Tuan et al.,  $\underline{Gene}$ , 90: 21 (1990)) and

Sac. erythraea strain NRRL2338 have been deposited in the culture collection of the Agricultural Research Laboratories, Peoria, IL and are available under the accession numbers NRRL XXXX, respectively. Staphylococcus aureus ThR (thiostrepton resistant) is obtained by plating 10<sup>8</sup> cells of S. aureus on agar medium containing 10 mg/ml thiostrepton and picking a survivor after 48 hr growth at 37°C. Thiostrepton is obtained from Squibb-Bristol Myers, New Brunswick, NJ. All other chemical and reagents are from standard commercial sources unless specified otherwise.

10

35

### **DNA Manipulations**

Standard conditions (Maniatis et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982) are employed for restriction endonuclease digestion, agarose gel-electrophoresis, nick translation of DNA to make <sup>32</sup>P-labeled probes, DNA ligation, and transformation of E. coli employing selection for ampicillin resistance (ApR) on LB agar plates. Plasmid DNA is isolated from minipreps of E. coli transformants by the boiling method (Maniatis et al., 1982, supra). DNA fragments are recovered from low melting agarose gels using the method of Langridge et al., 1980. Total DNA from 20 Sac. erythraea strains is prepared according to described procedures (Hopwood et al., Genetic Manipulation of Streptomyces, A Laboratory Manual, John Innes Foundation, Norwich, U.K., 1985). DNA is transferred from agarose gels onto Hybond N following the manufacturer's instructions. Hybridizations are performed in sealed bags 25 containing 10-20 ml of [1xNET (20xNET = 3 M NaCl, 0.3 M TrisHCl, 20 mM Na<sub>2</sub>EDTA, pH 8.0), 5XDenhardt's solution (Maniatis et al., 1982, supra), 0.2 mg/ml denatured calf thymus DNA, 0.2% SDS, and 0.5-2x10<sup>7</sup> cpm of the nick-translated probe] for 16-20 hr at 65 °C. Filters are washed three times in 1xNET/0.1% SDS for 20 min each at room temperature, and once in 30 0.05xNET/0.1% SDS for 20 min at 70 °C. Filters are reused as described (Donadio et al., 1990).

### Amplification of DNA fragments

Synthetic deoxyoligonucleotides are synthesized on an ABI Model 380A synthesizer (Applied Biosystems, Foster City, CA) following the manufacturer's recommendations. Amplification of DNA fragments is performed by the polymerase chain reaction (PCR) employing a Coy

10

15

20

25

30

35

Z,

Ė

thermocycler. Reactions contain 100 pmol c such primer, 1 µg of template DNA (cosmid pS1 carrying the *eryA* segment from Sac. erythraea strain NRRL 2338), and 2.5 units of Thermus aquaticus DNA polymerase in a 100 ml volume of PCR buffer [50 mM KCl, 10 mM TrisHCl (pH 8.0) 2 mM MgCl<sub>2</sub>, 0.01% gelatin) containing 200 mM of the 4 dNTPs. The above reagents are from Perkin Elmer Cetus, Norwalk, CT. The reaction mixture is overlaid with a drop of paraffin oil and subjected to 30-50 cycles. Each cycle consists of one 94 °C, one 55 °C and one 72 °C period, each of the duration of 3 min. The progress of the amplification is monitored by agarose gel-electrophoresis. The PCR primers described in the examples below are derived from the nucleotide sequence of *eryA* of FIG. 2.

### Gene replacement and gene conversion

Protoplasts of Sac. erythraea strains are prepared and transformed with miniprep DNA isolated from <u>E</u>. <u>coli</u> according to published procedures (Yamamoto et al., 1986). Integrative transformants, in the case of pWHM3 derivatives, are selected after one round of non-selective growth of the primary Th<sup>R</sup> transformants as described by Weber et. al, <u>Gene</u>, 68: 173 (1988). Loss of the Th<sup>R</sup> phenotype is monitored by plating serial dilutions of a Th<sup>R</sup> integrant on non-selective medium, followed by replica-plating on thiostrepton-containing medium. Th<sup>S</sup> (thiostrepton-sensitive) colonies arise at a frequency of 10<sup>-2</sup> (Donadio et al., 1990). The retention of the mutant allele is established by Southern hybridization of a few Th<sup>S</sup> colonies.

A few hundred Th<sup>R</sup> colonies obtained by transformation of an *eryA* strain with pWHM4 derivatives are screened for antibiotic production by the agar-plug assay employing *Staphylococcus aureus* as Th<sup>R</sup> organism as described (Tuan et al., Gene, 90: 21 (1990)). The frequency of gene conversion between a 5 kb segment of homologous sequence and a strain carrying a small deletion is >25% (Tuan et al., Gene, 90: 21 (1990)). Colonies found to produce antibiotic activity are inoculated in SGGP (Yamamoto et al., 1986), protoplasts are prepared, and the regenerated protoplasts are scored for loss of the plasmid by replica-plating on non-selective medium. Th<sup>S</sup> colonies are then rechecked for antibiotic production, and six producers are analyzed on Southern blots.

### **Fermentation**

Sac. erythraea cells are inoculated into 100 ml SCM medium (1.5% soluble starch, 2.0% Soytone [Difco], 0.15% Yeast Extract [Difco], 0.01% CaCl<sub>2</sub>) and allowed to grow at 32°C for 3 to 6 days. The entire culture is then inoculated into 10 liters of fresh SCM medium. The fermenter is operated for a period of 7 days at 32°C maintaining constant aeration and pH at 7.0. After fermentation is complete, the cells are removed by centrifugation at 4°C and the fermentation beer is kept in the cold until further use.

The present invention will now be illustrated, but is not intended to be limited, by the following examples:

## Example 1 Construction of plasmid pABX9

15

10

The 9.6 kb <u>BamHI-XhoI</u> segment comprised between sequence coordinates 21.96 and 31.52 was isolated from cosmid pS1 and ligated to <u>SalI</u>-digested pUC19 DNA. The resulting mixture contained the desired plasmid pABX9.

20

## Example 2 Construction of E. coli K12 DH5α/pABX9

Approximately 10 ng of plasmid pABX9, prepared as described in Example 1, were transformed into <u>E</u>. <u>coli</u> K12 DH5α and a few of the resulting white ApR colonies that appeared on the LB-agar plates containing X-gal and ampicillin were analyzed for their plasmid content. One colony was found to carry pABX9, as verified by the observation of fragments of 3.93, 3.39, 2.01, 1.56, 0.87, and 0.48 kb in size upon agarose gel electrophoresis after <u>SmaI</u> digestion of the plasmid.

## Example 3 Construction of plasmid pABX9DN

Plasmid pABX9, isolated from <u>E</u>. <u>coli</u> K12 DH5 $\alpha$ /pABX9, was digested with <u>Nco</u>I and then treated with T4 DNA ligase. The resulting mixture contained the desired plasmid pABX9DN.

25

30

35

\*

ń

£

## Example 4 Construction of E. coli K12 DH5a/pABX9DN

Approximately 10 ng of plasmid pABX9DN, prepared as described in Example 3, were transformed into <u>E</u>. <u>coli</u> K12 DH5α and a few of the resulting white Ap<sup>R</sup> colonies that appeared on the LB-agar plates containing X-gal and ampicillin were analyzed for their plasmid content. Colonies carrying pABX9DN exhibited a single <u>Ncol</u> fragment of 11.5 kb visible by agarose gel electrophoresis, confirming that the 813 bp <u>Ncol</u> - <u>Ncol</u> fragment from pABX9 has been deleted in pABX9DN.

## Example 5 Construction of plasmid pABX95DN

Plasmid pABX95DN was digested with <u>EcoRI</u> and <u>HindIII</u> and ligated to pWHM3 digested with the same two enzymes. The resulting mixture contained the desired plasmid pABX95DN.

## Example 6 Construction of E. coli K12 DH5α/pABX95DN

Approximately 10 ng of plasmid pABX95DN, prepared as described in Example 5, were transformed into <u>E</u>. <u>coli</u> K12 DH5a and a few of the resulting white Ap<sup>R</sup> colonies that appeared on the LB-agar plates containing X-gal and ampicillin were analyzed for their plasmid content. Colonies carrying pABX95DN exhibited fragments of 8.8 and 7.2 kb visible in agarose gels after <u>Eco</u>RI and <u>HindIII</u> digestion.

# Example 7 Construction of Sac. erythraea AKR5 carrying the eryAKR5 allele by gene replacement

Approximately 1 mg of plasmid pABX95DN, isolated from <u>E</u>. <u>coli</u> K12 DH5α/pABX95DN, was transformed into *Sac. erythraea* NRRL 2338 and stable Th<sup>R</sup> colonies were isolated. Serial dilutions of one of these colonies were screened for the loss of the antibiotic resistance marker and total DNA from 5 Th<sup>S</sup> colonies as well as from untransformed *Sac. erythraea* NRRL 2338 was digested with <u>Ss</u>tI and analyzed by Southern

ž

5

hybridization employing the 0.8 kb <u>SalI</u> fragment between sequence coordinates 24.26 and 25.06 (from pABX9) as probe. Whereas NRRL 2338 showed one <u>SstI</u> band of 3.7 kb that hybridized to the probe, samples from four of the Th<sup>S</sup> strains exhibited a <u>SstI</u>-hybridizing band of 6.1 kb indicating the presence of the mutant allele. One of these colonies was kept and designated strain AKR5. It carries a deletion of 813 bp in the KR5 segment of *eryA* and is designated the <u>eryAKR5</u> allele.

### Example 8

### 10 <u>Isolation, purification and properties of 5-oxo-5,6-dideoxy-3-a-mycarosyl</u> erythronolide B from Sac. erythraea AKR5

A 10-liter fermentation of Sac. erythrea AKR5 carrying the eryAKR5 allele in a Biolafitte fermentor using SNC Media. The fermentor was inoculated with 100 ml of a 3 day old seed. The pO2 was initially 80 ppm 15 and the temperature was maintained at 32°C. The pH was controlled to  $7.0 \pm 0.2$  by addition of propionic acid or potassium hydroxide as needed. At harvest (3 days), the whole broth was extracted three times with 4-liter portions of ethylacetate. The combined extracts were concentrated under reduced pressure and the residue was chromatographed on a column (50  $\times$ 20 5 cm) of silica gel packed and loaded in toluene and eluted with a stepwise gradient of increasing concentration of isopropanol in toluene. Fractions were analyzed by TLC and spots were detected by spraying with anisaldehyde sulfuric acid spray reagent and heating. A major component giving blue colored spots eluted with approximately 7% isopropanol. 25 Fractions containing this material were combined and concentrated to a residue (800 mg). This was further chromatographed on a column (100  $\times$ 3 cm) of Sephadex LH-20 in chloroform-heptane-ethanol, 10:10:1, v/v/v. Fractions were analyzed as above, early fractions (9-13) yielded 5,6-dideoxy-3-a-mycarosyl-5-oxoerythronolide B (45 mg) which was crystallized from 30 heptane/ethylacetate mixture to mp 163-164 °C.

\*

Ġ

### CMR spectrum in CDCl3 (ppm downfield from TMS)

9.9       38.7       76.2         9.9       40.4       76.4         10.4       40.7       80.4         14.5       43.3       100.4         15.2       45.8       175.8         17.1       46.8       210.8	8.6	37.9	70.0
10.4       40.7       80.4         14.5       43.3       100.4         15.2       45.8       175.8         17.1       46.8       210.8         17.7       48.9       217.3         25.3       66.5		38.7	76.2
14.5       43.3       100.4         15.2       45.8       175.8         17.1       46.8       210.8         17.7       48.9       217.3         25.3       66.5	9.9	40.4	76.4
15.2     45.8     175.8       17.1     46.8     210.8       17.7     48.9     217.3       25.3     66.5	10.4	40.7	80.4
17.1     46.8     210.8       17.7     48.9     217.3       25.3     66.5	14.5	43.3	100.4
17.1 17.7 48.9 25.3 66.5	15.2	45.8	175.8
25.3 66.5	17.1	46.8	210.8
	17.7	48.9	217.7
25.5 69.4	25.3	66.5	
	25.5	69.4	

Structure was determined by single crystal X-ray diffraction.

Later fractions (15-17) yielded 5,6-dideoxy-5-oxoerythronolide B (10 mg) and still later fractions yielded 5,6-dideoxy-6,6a-epoxy-5-oxoerythronolide B (2.8 mg).

### 10 <u>Example 9</u> Construction of plasmid pALeryAKR2

contains the desired plasmid pALeryAKR2.

The 1.3 kb DNA segment comprised between coordinates 8.63-9.93

(fragment 1) is amplified by PCR employing two oligodeoxynucleotides, 1a

15 (5'-GGGAGCATGCTCTCGGTGCGCGGGCGGCGCGC-3') and 1b (5'GCCCTGCAGCGCGTACTCCGAGGTGGCGGT-3'). Similarly, the 1.3 kb

DNA segment between coordinates 9.99-11.26 (fragment 2) is PCRamplified employing primers 2a (5'TGGTCTGCAGGCGAGGCCGGACACCGAGG-3') and 2b (5'
20 GGAAGAAGTCAAAGTTCCTCGGTCCCTTCT-3'). After digestion with

SphI + PstI (fragment 1) and PstI + EcoRI (fragment 2), the two fragments
are ligated to EcoRI + SphI-digested pWHM3. The resultant mixture

## Example 10 Construction of E. coli K12 DH5a/pALeryAKR2

Approximately 10 ng of plasmid pALeryAKR2, prepared as

described in Example 9, are transformed into Ε. coli K12 DH5α, and a few
of the resulting white ApR colonies that appear on the LB-agar plates
containing X-gal and ampicillin are analyzed for their plasmid content.
The identity of plasmid pALeryA2KR2, 9.8 kb in size, and carrying a 2.6 kb
EcoRI-SphI insert with an internal PstI site, is verified by SalI digestion
(fragments at 2.91, 2.21, 1.61, 1.42, 1.08, 0.29, 0.12 and 0.10 kb are released,
visible by agarose gel electrophoresis). pALeryAKR2 contains an in-frame
deletion of 102 base pairs of the corresponding segment of the wild type
eryA chromosomal DNA. The cloned segment in pALeryAKR2 is
designated the eryAKR2 allele.

15

# Example 11 Construction of Sac. erythraea AKR2 carrying the eryAKR2 allele by gene replacement

2

\*

T

Approximately 1 mg of plasmid pALeryAKR2, isolated from E. coli 20 K12 DH5α/pALeryAKR2, is transformed into Sac. erythraea protoplasts and stable ThR colonies are isolated. Serial dilutions of one of these colonies are screened for loss of the antibiotic resistance marker, and six Th<sup>S</sup> colonies are analyzed for their genotype by Southern hybridization. Total DNA from the six ThS colonies and from untransformed Sac. 25 erythraea NRRL2338 is digested with PstI and with SalI and is then examined by Southern hybridization using the 2.6 kb EcoRI-SphI insert from pALeryAKR2 as probe. Whereas NRRL2338 contains a 39 kb PstI hybridizing band, colonies in which the mutation in KR2 has been introduced (strain AKR2) exhibit two bands of approximately equal 30 intensity, one at 27 kb and the other at 12 kb. The SalI digest, with bands at 1.04, 0.75, 0.29, 0.12 and 0.10 kb common to NRRL2338 and AKR2, but with the 1.16 kb band in NRRL2338 replaced by the 1.06 kb band in AKR2, confirms that the only change introduced into strain AKR2 is the deletion of the 102 bp segment from KR2, resulting in a strain carrying the 35 eryAKR2 allele.

â

Ž

### Example 12

### Isolation and purification of 11-deoxy-11-oxoerythromycin A

The fermentation beer of strain AKR2, cooled to 4°C is adjusted to pH 8.0 and is extracted sequentially with three equal volumes of methylene chloride. The combined methylene extracts are concentrated to an oily residue and partitioned between heptane and methanol. The methanol layer is removed, washed once with heptane and concentrated to a residue. The residue is digested in methylene chloride and washed once with potassium phosphate buffer pH 7.8 and once with water. The 10 methylene chloride layer is concentrated to a residue and digested in the lower phase (1:1:1, v/v/v) of a carbon tetrachloride; methanol; aqueous phosphate buffer (0.05 M, pH 7.0) system and chromatographed on an Ito Coil Planet Centrifuge in the same system. Fractions containing the desired 11-oxo-11-deoxyerythromycin A were combined, concentrated, 15 digested in methylene chloride, washed well with water and concentrated on rotary evaporator under reduced pressure to yield 11-deoxy-11oxoerythromycin A as an off-white solid froth. Its identity is confirmed by comparison with antibiotic L53-18A. 11-Deoxy-11-oxoerythromycin A is dissolved in tetrahydrofuran and the solution is diluted with an equal 20 volume of water. This is then acidified to pH 4.0 and allowed to stand at room temperature for 4 hours. The pH is adjusted to 9.0 and the solution is diluted with an equal volume of water and extracted with two volumes of methylene chloride. The combined methylene chloride extracts were evaporated to dryness under reduced pressure to yield antibiotic L53-18A 25 as a white solid.

## Example 13 Construction of plasmid pALeryADH4

30

35

Primers 3a (GCGCGAGCTCGACGACCAGGGCGGCATGGT) and 3b (GGTGGCATGCTGCGACCACTGCGCGTCGGC) are used to PCR-amplify the 1.05 kb *eryA* segment of the <u>Sac. erythraea</u> chromosome between sequence coordinates 18.47-20.07 (fragment 3), and primers 4a (AGCTGCATGCTCTGGACTGGGGACGGCTAG) and 4b (CGCGGGATCCCAGCTCCCACGCCGATACCG) are used to amplify the 1.35 kb segment between sequence coordinates 20.58-21.96 (fragment 4) as described in Example 1. Fragment 3 and 4, after digestion with <u>SstI + SphI</u>

Ê

F

20

25

30

35

and with <u>SphI + BamHI</u>, respectively, are ligated to <u>SstI -, BamHI</u>-digested pWHM3. The resulting ligation mixture contains the desired plasmid pALeryADH4.

## Example 14 Construction of E. coli K12 DH5a/pALeryADH4

Approximately 10 ng of pALeryADH4, prepared as described in Example 13, are transformed transformed into Ε. coli K12 DH5α, and a few of the resulting white ApR colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryADH4, 9.6 kb in size, is verified by SphI + EcoRI digestion (fragments at 7.2, 1.35 and 1.05 kb are released). pALeryADH4 carries a 498 base pair in-frame deletion of the corresponding segment of the wild type eryA DNA. The cloned segment in pALeryADH4 is designated the eryADH4 allele.

# Example 15 Construction of Sac. erythraea ADH4 carrying the eryADH4 allele by gene replacement

Approximately 1 mg of plasmid pALeryADH4, isolated from E. coli K12 DH5α/pALeryADH4, is used for transformation into Sac. erythraea protoplasts and stable ThR colonies are isolated. Serial dilutions of one of these colonies are screened for loss of the antibiotic resistance marker, and six Th<sup>S</sup> colonies are analyzed for their genotype by Southern hybridization. Total DNA from the six Th<sup>S</sup> colonies and from untransformed Sac. erythraea NRRL2338 is digested with SphI and with SstI and examined by Southern hybridization using the 2.4 kb SstI-BamHI insert from pALeryADH4 as probe. Strains in which the wild type allele has been replaced by the mutated copy show two SphI bands, one at 13.5 kb and the other at 12.4 kb, whereas the wild type strain exhibits a single band at 26 kb. The SstI pattern, with the 2.9 kb band from NRRL2338 being replaced in ADH4 by a 2.5 kb band, confirms that the 487 bp deletion created in plasmid pALeryADH4 has been transferred into the chromosome of ADH4. Strains that carry the eryADH4 allele in place of the wild type sequence are designated Sac. erythraea ADH4.

3

5

矢

•1

### Example 16

## Isolation and characterization of 7-hydroxyerythromycin A and 6-deoxy-7-hydroxyerythromycin A

The fermentation beer of strain ADH4 is cooled to 4°C and the pH is 5 adjusted to 5.0. The mixture is extracted once with an equal volume of methylene chloride. The pH of the aqueous layer is readjusted to 9.0 and two further methylene chloride extracts are carried out. These two extracts are combined, washed with water and concentrated to a residue. This is digested in 10 ml of the upper phase of a (3:7:5, v/v/v) mixture of hexane, 10 ethylacetate, aqueous phosphate buffer (0.05 M, pH 7.5) and chromatographed on an Ito Coil Planet Centrifuge in the same system. Fractions containing the desired 7-hydroxyerythromycin were combined, concentrated, and partitioned between methylene chloride and dilute (pH 9.5) ammonium hydroxide solution. Fractions containing the desired 6deoxy-7-hydroxyerythromycin were combined, concentrated, and partitioned between methylene chloride and dilute (pH 9.5) ammonium hydroxide solution. The methylene chloride layers are washed with water and then concentrated to yield the desired 7-hydroxyerythromycin A and 6-deoxy-7-hydroxyerythromycin A as white foams. 20

## Example 17 Construction of plasmid pALeryAKS1

- The 1.4 kb segment of eryA, between sequence coordinates 1.11-2.54 (fragment 5) and the 1.5 kb segment between sequence coordinates 2.88-4.37 (fragment 6) are PCR-amplified using primers 5a (TGCAGAATTCGCTGGCCGCGCTCTGGCGCT) and 5b (GAGAGCTGCAGCATGAGCCGCTGCTGCGGG), and 6a
- (CATGCTGCAGGACTTCAGCCGGATGAACTC) and 6b (GAGGAAGCTTCCAGCCGGTCCAGTTCGTCC), respectively, as described in Example 9. After digestion with <a href="EcoRI + PstI">EcoRI + PstI</a> (fragment 5) and <a href="PstI">PstI</a> (fragment 5) and <a href="PstI">PstI</a> (fragment 5) and <a href="PstI">PstI</a> + <a href="HindIII">HindIII</a> + <a href="HindIII">HindIII</a> <a href="HindIII"
- 35 pALeryAKS1.

## Example 18 Construction of E. coli K12 DH5a/pALeryAKS1

3

1

3

•

Approximately 10 ng of pALeryAKS1, prepared as described in

5 Example 17, are transformed into <u>E</u>. <u>coli</u> K12 DH5α, and a few of the resulting white ApR colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAKS1, 10.1 kb in size, is verified by digestion with <u>PstI</u> + <u>HindIII</u> (fragments of 8.6 and 1.5 kb are observed by agarose gel electrophoresis) and with <u>SalI</u> (fragments of 2.93, 2.21, 1.42, 1.37, 0.86, 0.54, 0.27, 0.14, 0.13, and 0.10 kb are observed). pALeryAKS1 carries an in-frame deletion of 282 base pairs of the corresponding wild type *eryA* DNA. The cloned insert in plasmid pALeryAKS1 is designated the <u>eryAKS1</u> allele.

# Example 19 Construction of Sac. erythraea AKS1 carrying the eryAKS1 allele by gene replacement

Approximately 1 mg of plasmid pALeryAKS1, isolated from E. coli K12 DH5α/pALeryAKS1, is used for transformation into Sac. erythraea 20 protoplasts and stable  $\mathsf{Th}^R$  colonies are isolated. Serial dilutions of one of these colonies are screened for loss of the antibiotic resistance marker, and six ThS colonies are analyzed for their genotype by Southern hybridization. Total DNA from the six ThS colonies and from untransformed Sac. erythraea NRRL2338 is digested with PstI and with 25 Smal and examined in Southern hybridization employing the 2.9 kb EcoRI-HindII insert from pALeryAKS1 as probe. Colonies in which the wild type allele has been replaced by the mutated copy (strain AKS1) show two PstI bands, one at 34.5 and the other at 4.4 kb, whereas the wild type strain exhibits a single band at 39 kb. The SmaI pattern, with the 2.9 kb 30 band from NRRL2338 being replaced in AKS1 by a 2.6 kb band, confirms that the 282 bp created in plasmid pALeryAKS1 has been transferred into strain AKS1. Strains that carry the eryAKS1 allele are designated Sac. erythraea AKS1.

15

20

30

35

5

3

### Example 20

### Synthesis of (2S,3R,4S,5S)3,5-dihydroxy-2,4-dimethylhexanoic acid n-butyl thioester

A convenient source of this compound in chiral purity is the antibiotic oleandomycin. Oleandomycin (5 g) is dissolved in an aprotic solvent such as toluene and treated with diazabicyclo[5.4.0]undecene-5 (1 g) and heated for one hour. The resulting solution is poured into iced water, agitated well and the organic layer is drawn off and concentrated to a residue. The residue is digested in methylene chloride and treated exhaustively with a solution of ozone. The resulting ozonide is oxidatively decomposed with dilute hydrogen peroxide in sufficient aqueous ethanol to yield a monophasic mixture. This is further diluted with water and made 0.1 N with sodium hydroxide. The mixture is warmed for one hour at 70°C and then cooled before being acidified to 15 pH 2.5 with dilute sulfuric acid. The mixture is then exhaustively extracted with methylene chloride. The combined extracts are concentrated to an oily residue and the desired lactone is recovered by chromatography on silica gel eluted with a gradient of tolueneisopropanol.

The  $\delta$ -lactone is converted to the butyl thioester before feeding to Sac. erythrea AKS1 by refluxing with n-butylthiol in the presence of a catalytic amount of triethylamine.

#### Example 21 25 Isolation of (14S,15S)14(1-hydroxyethyl)erythromycin A

The fermentation broth of AKS1 is cooled to 4°C and adjusted to pH 4.0 and extracted once with methylene chloride. The aqueous layer is readjusted to pH 9.0 and extracted twice with methylene chloride and the combined extracts are concentrated to a solid residue. This is digested in methanol and chromatographed over a column of Sephadex LH-20 in methanol. Fractions are tested for bioactivity against a sensitive organism, such as Staphylococcus aureus ThR, and active fractions are combined. The combined fractions are concentrated and the residue is digested in 10 ml of the upper phase of a solvent system consisting of n-heptane, benzene, acetone, isopropanol, 0.05 M, pH 7.0 aqueous phosphate buffer (5:10:3:2:5, v/v/v/v), and chromatographed on an Ito Coil Planet

WO 93/13663

22

PCT/US92/00427

ど

5

3

Ē

Centrifuge in the same system. Active fractions are combined, concentrated and partitioned between methylene chloride and dilute ammonium hydroxide (pH 9.0). The methylene chloride layer is separated and concentrated to yield the desired product as a white foam.

5

20

## Example 22 Construction of plasmid pALeryAKS2

Primers 7a (CGCCCGAATTCGAGGCGCTGGGCGCCCGGAC) and 7b

(CCACCTGCAGCGCGGGACCTTCCAGCCCC), and primers 8a

(GTGGGTCGCTGCAGACGGTGACTGCGG) and 8b

(GGTCAAGCTTCGTCGGCGAGCAGCTTCTC) are used to PCR-amplify

the 1.45 kb eryA segment between sequence coordinates 5.71-7.16

(fragment 7) and the 1.5 kb eryA segment between sequence coordinates

7.22-8.70 (fragment 8), respectively. After digestion with EcoRI + PstI

(fragment 7) and with PstI + HindIII (fragment 8), the two fragments are ligated to pWHM3 cut with EcoRI + HindIII. The resulting mixture contains the desired plasmid pALeryAKS2.

## Example 23 Construction of E. coli K12 DH5a/pALeryAKS2

Approximately 10 ng of pALeryAKS2, prepared as described in Example 22, are transformed into E. coli K12 DH5\(\alpha\_z\) and a few of the 2.5 resulting white ApR colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAKS2, 10.1 kb in size, is verified by digestion with PstI + HindIII (fragments of 8.6 and 1.5 kb are observed by agarose gel electrophoresis) and with SstII (fragments of 4.0, 2.3, 2.0, 0.72, 0.43, 0.40, 0.20, 0.18, 0.13 and 0.11 kb observed). Plasmid pALeryAKS2 carries an inframe deletion of 60 base pairs of the corresponding wild type eryA DNA. This deletion removes the active site cysteine from KS2. The cloned insert in plasmid pALeryAKS2 is designated the eryAKS2 allele.

녕

### Example 24

## Construction of Sac. erythraea AKS2 carrying the eryAKS2 allele by gene replacement

Approximately 1 mg of plasmid pALeryAKS2, isolated from E. coli 5 K12 DH5 $\alpha$ /pALeryAKS2, is used for transformation into Sac. erythraea protoplasts and stable ThR colonies are isolated. Serial dilutions of one of these colonies are screened for loss of the antibiotic resistance marker, and six ThS colonies are analyzed for their genotype by Southern hybridization. Total DNA from the six ThS colonies and from 10 untransformed Sac. erythraea NRRL2338 is digested with PstI and with SstII and examined in Southern hybridization employing the 2.9 kb EcoRI-HindII insert from pALeryAKS2 as probe. Colonies in which the wild type allele has been replaced by the mutated copy (strain AKS2) show two PstI bands, one at 34.5 and the other at 4.4 kb, whereas the wild type strain 15 exhibits a single band at 39 kb. The SstII pattern, with the 0.78 kb band from NRRL2338 being replaced in AKS2 by a 0.72 kb band, confirms that the 60 bp created in plasmid pALeryAKS2 has been transferred into strain AKS2. Strains that carry the eryAKS2 allele are designated Sac. erythraea AKS2. 20

### Example 25

## Synthesis of (2R,3R,4S,5R)2,4-dimethyl-3-fluoro-5-hydroxyhexanoic acid n-butyl thioester

25

30

35

**李** 

(2R,3S,4S,5R)3,5-Dihydroxy-2,4-dimethylhexanoic acid-δ-lactone (1 g) from Example 20 is digested in 10 ml of pyridine and treated with p-toluenesulfonyl chloride (1.3 g) and allowed to stand at room temperature overnight. The mixture is poured into iced water and extracted with methylene chloride and the methylene chloride is concentrated to the crude sulfonate ester. This is digested in acetonitrile (100 ml) and heated under reflux after the addition of tetrabutylammonium fluoride (1.75 g). After 6 hours the mixture is cooled, poured over iced water (300 ml) and extracted three times with 200 ml portions of methylene chloride. The combined methylene chloride extracts were concentrated and the residue was chromatographed on a column of silica gel eluted with a stepwise gradient of isopropanol (0 to 50%) in toluene. Fractions containing (2R,3R,4S,5R)2,4-dimethyl-3-fluoro-5-hydroxyhexanoic acid\_d-lactone were

**†** 

क्र

Ė

4

combined and concentrated to a white solid. The lactone is then converted to the n-butyl thiolester by refluxing in n-butyl thiol with a catalytic amount of triethylamine. Solvent is removed and the residue is digested in DMSO before feeding to fermentations of <u>Sac. erythraea</u> AKS2.

5

25

30

## Example 26 Isolation and purification of 11-epifluoro-15-norerythromycin A

The fermentation broth of strain AKS2 is cooled to 4°C and adjusted to pH 4.0 and extracted once with ethylacetate. The aqueous layer is 10 adjusted to pH 9.0 and extracted twice with methylene chloride and the combined extracts are concentrated to a white solid. This is chromatographed over a column of Sephadex LH-20 in a mixture of heptane, chloroform, ethanol (10:10:1, v/v/v) and fractions containing the desired product are combined and concentrated to a solid residue. This is 15 further purified by countercurrent chromatography on an Ito Coil Planet Centrifuge on a system composed of carbon tetrachloride; methanol; 0.05 M; pH 7.0 aqueous potassium phosphate buffer (1:1:1, v/v/v). Fractions containing the desired 11-epifluoro-15-norerythromycin were combined, and concentrated to a residue. This was digested in methylene chloride 20 and dilute (pH 9.5) ammonium hydroxide and the methylene chloride layer was separated, washed with water and concentrated to yield the desired 11-epifluoro-15-norerythromycin A as white solid.

## Example 27 Construction of plasmid pALeryAM4.1

Primers 9a (GCGCCGAATTCTCGAGACGCGTGGGAGGCA) and 9b (TTGCGGTACCAGTAGGAGGCGTCCATCGCG) are employed to PCR-amplify the 2.0 kb *eryA* segment between sequence coordinates 17.35-19.38 (fragment 9). After digestion with <u>EcoRI + KpnI</u>, fragment 9 is ligated to pUC19 cut with the same two enzymes The resulting mixture contains the desired plasmid pALeryAM4.1.

÷ .

## Example 28 Construction of E. coli K12 DH5a/pALeryAM4.1

Approximately 10 ng of pALeryAM4.1, prepared as described in

5 Example 27, are transformed into <u>E</u>. <u>coli</u> K12 DH5a, and a few of the
resulting white Ap<sup>R</sup> colonies that appear on the LB-agar plates containing
X-gal and ampicillin are analyzed for their plasmid content. The identity
of plasmid pALeryAM4.1, 4.7 kb in size, is verified by digestion with <u>Sall</u>
(fragments of 2.8, 0.85, 0.53, 0.27 and 0.22 kb are observed by agarose gel
electrophoresis).

## Example 29 Construction of plasmid pALeryAM4.2

Primers 10a (GCTGGGATCCCGCGGGGGTTGCAGCAC) and 10b (CGGAACTCGGTGAGCATGCCGGGACTGCTC) are used to PCR-amplify the 2.1 kb *eryA* segment between sequence coordinates 21.94-24.00 (fragment 10). The 2.6 kb fragment KpnI(96)-BamHI(102) from cosmid clone pR1, and fragment 10 cut with BamHI + SphI, are ligated to pALeryAM4.1 cut with KpnI + SphI. The resulting mixture contains the desired plasmid pALeryAM4.2.

### Example 30 Construction of E. coli K12 DH5a/pALeryAM4.2

25

30

Approximately 10 ng of pALeryAM4.2, prepared as described in Example 29, are transformed into <u>E. coli</u> K12 DH5a, and a few of the resulting white Ap<sup>R</sup> colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM4.2, 9.3 kb in size, is verified by digestion with <u>XhoI</u> + <u>SphI</u> (to ensure that the entire 6.65 kb insert is released) and with <u>SalI</u>, with fragments of 2.8, 1.82, 1.09, 0.94, 0.85, 0.75, 0.45, 0.27, 0.22 and 0.13 kb are observed by agarose gel electrophoresis).

## Example 31 Construction of plasmid pALeryAM1

The 2.9 kb <u>SmaI(4)-SmaI(20)</u> fragment from cosmid clone pR1 is ligated to pUC12 cut with <u>SmaI</u>. The resulting mixture contains plasmid pALeryAM1.

## Example 32 Construction of E. coli K12 DH5αa/pALeryAM1

3

ğ

Ť

\*

Approximately 10 ng of pALeryAM1, prepared as described in Example 31, are transformed into E. coli K12 DH5α, and a few of the resulting white ApR colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM1, 5.6 kb in size, is verified by digestion with Smal (the 2.9 kb insert is realeased) and with SphI, with release of one 4.4 and one 1.07 kb bands. Both orientations of the insert in plasmid pALeryAM1 are useful.

## Example 33 Construction of plasmid pALeryAM4.3

20

30

35

Plasmid pALeryAM1 is cut with XhoI to completion, partially with SphI, and the resulting 5.25 kb band, isolated from an agarose gel, is ligated to the 6.65 kb insert released from pALeryAM4.2 by XhoI + SphI digestion The resulting mixture contains the desired plasmid pALeryAM4.3.

## Example 34 Construction of E. coli K12 DH5a/pALeryAM4.3

Approximately 10 ng of pALeryAM4.3, prepared as described in Example 33, are transformed into  $\underline{E}$ . coli K12 DH5 $\alpha_z$  and a few of the resulting white  $Ap^R$  colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM4.1, 11.9 kb in size, is verified by  $\underline{XhoI} + \underline{SphI}$  digestion (fragments of 6.65 and 5.25 kb are visible by agarose gelelectrophoresis). Plasmid pALeryAM4.3 carries the entire eryA module 4

10

3

inserted into the KS region of module 1. The cloned insert in pALeryAM4.3 is degnated the <u>eryAM412</u> allele.

## Example 35 Construction of plasmid pALeryAM4.4

Plasmid pALeryAM4.3 is cut with <u>Eco</u>RI + <u>Hind</u>III, and the resulting 9.2 kb band, recovered from an agarose gel, is ligated to pWHM4 cut with the same two enzymes. The resulting mixture contains the desired plasmid pALeryAM4.4.

## Example 36 Construction of E. coli K12 DH5\alpha/pALeryAM4.4

Approximately 10 ng of pALeryAM4.4, prepared as described in Example 35, are transformed into <u>E</u>. coli K12 DH5α, and a few of the resulting white Ap<sup>R</sup> colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM4.1, 16.5 kb in size, is verified by <u>Eco</u>RI + <u>HindIII</u> digestion, with fragments of 9.2 and 7.3 kb released. Plasmid pALeryAM4.4 carries the <u>eryAM412</u> allele on the <u>Sac. erythraea</u> multicopy vector pWHM4.

### Example 37

### 25 <u>Construction of Sac. erythraea AM412 carrying the eryAM412 allele by</u> gene conversion

Approximately 1 mg of plasmid pALeryAM4.4, isolated from E. coli
K12 DH5\(aappa/PALeryAM4.4\), is used for transformation into Sac. erythraea
30 strain AKS1 protoplasts. A few hundred transformants are screened for antibiotic production by the agar-plug assay, and one of the colonies found to produce antimicrobial activity is cured of pALeryAM4.4 by protoplast formation and regeneration as described in General Methods. Total DNA from six antibiotic-producing, ThS colonies (strain AM412)and from strain
35 AKS1 is digested with SphI and with XhoI and the resulting Southern blot is hybridized first to the 2.9 kb insert from pALeryAM1, and then to the 2.9 kb SstI(95)-SstI(101) fragment from plasmid pALeryAM4.2. With the first probe, the SphI band at 0.8 kb in strain AKS1 is seen to be replaced by a 7.5

WO 93/13663 PCT/US92/00427

28

kb band in strain AM412, whereas the other two bands at 2.4 kb and 5.2 kb are unaffected. In the XhoI digest, the AKS1 band at 2.9 kb is replaced by a 9.6 kb band in AM412, with the other band at 5.2 kb conserved in both strains. Using the SstI(95)-SstI(101) fragment as probe, strain AKS1 exhibits one band at 25.5 kb and one at 17.9 kb in the SphI and XhoI digests, respectively, whereas, in addition to these bands, strain AM412 shows one SphI band at 7.5 kb and one XhoI band at 9.6 kb. In this way, it is established that the eryAKS1 allele has been converted into the eryAM412 allele in strain AM412.

贵

OF:

ķ

10

• •

## Example 38 Isolation and purification of 14-(1-propyl)erythromycin A

twice with equal volumes of methylene chloride. The combined extracts are washed once with water and concentrated to an oily residue. This is partitioned in a heptane methanol water (5:5:1, v/v/v) system and the lower layer is washed once with heptane and then concentrated to a semisolid residue. This is digested in methanol and chromatographed over a column of Sephadex LH-20 in methanol. Fractions are tested for bioactivity in an agar diffusion assay on plates seeded with the macrolide-sensitive strain <a href="Staphylococcus aureus ThR">Staphylococcus aureus ThR</a>. Active fractions are combined and further purified by chromatography over silica gel a chloroform:methanol gradient containing 0.1% triethylamine. Fractions containing the desired 14-(1-propyl)erythromycin A are combined and concentrated to yield the product as a white solid.

## Example 39 Construction of plasmid pALeryAM5.1

30

35

The 4.7 kb *eryA* fragment between sequence coordinates 23.65-28.36 (fragment 11) is PCR-amplified employing primers 11a (ATGCTCGAGATCTCGTGGGAGGCGCTGGA) and 11b (AGAACTCGGTGAGCATGCCCGGGCCCGCCA). Fragment 11, after digestion with <u>XhoI</u> + <u>SphI</u>, is ligated to the 5.25 kb fragment resulting from complete <u>XhoI</u> and partial <u>SphI</u> digestion of pALeryAM1, as in Example 33. The resulting mixture contains the desired plasmid pALeryAM5.1.

## Example 40 Construction of E. coli K12 DH5\alpha/pALeryAM5.1

Approximately 10 ng of pALeryAM5.1, prepared as described in Example 39, are transformed into <u>E</u>. <u>coli</u> K12 DH5α, and a few of the resulting white Ap<sup>R</sup> colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM5.1, 9.95 kb in size, is verified by <u>SphI + XhoI</u> digestion, with fragments of 5.25 and 4.7 kb released, and by <u>SmaI</u> digestion where fragments of 3.39, 2.68 and 1.94 (doublet) kb are observed. Plasmid pALeryAM5.1 carries the entire *eryA* module 5 inserted into the β-ketoacyl ACP synthase region of module1. The cloned insert in plasmid pALeryAM5.1 is designated the <u>eryA512</u> allele.

15

10

Ť

É

## Example 41 Construction of plasmid pALeryAM5.2

Plasmid pALeryAM5.1 is cut with <u>EcoRI + HindIII</u> and the resulting 6.3 kb fragment, recovered from an agarose gel, is ligated to pWHM4 cut with the same two enzymes. The resulting mixture contains the desired plasmid pALeryAM5.2.

## Example 42 Construction of E. coli K12 DH5α/pALeryAM5.2

30

25

Approximately 10 ng of pALeryAM5.2, prepared as described in Example 41, are transformed into E. coli K12 DH5α, and a few of the resulting white Ap<sup>R</sup> colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM5.2, 13.6 kb in size, is verified by digestion with EcoRI + HindIII, with fragments of 7.3 and 6.3 kb released. Plasmid pALeryAM5.2 contains the eryAM512 allele on the Sac. erythraea multicopy vector pWHM4.

35

Ë

₹

10

Ĵ

#### Example 43

## Construction of Sac. erythraea AM512 carrying the eryAM512 allele by gene conversion

Approximately 1 mg of plasmid pALeryAM5.2, isolated from E. coli 5 K12 DH5α/pALeryAM5.2, is used for transformation into Sac. erythraea strain AKS1 protoplasts. A few hundred transformants are screened for antibiotic production by the agar-plug assay, and one of the colonies found to produce antimicrobial activity is cured of pALeryAM5.2 by protoplast formation and regeneration as described in General Methods. Total DNA 10 from six antibiotic-producing, ThS colonies (strain AM512)and from strain AKS1 is digested with SphI and with XhoI and the resulting Southern blot is hybridized first to the 2.9 kb insert from pALeryAM1, and then to the 0.8 kb NcoI(119)-NcoI(123) fragment from plasmid pALeryAM5.1. With the first probe, the SphI band at 0.8 kb in strain AKS1 is replaced by a 5.5 kb band in strain AM512, whereas the other two bands at 2.4 kb and 5.2 kb are unaffected. In the XhoI digest, the AKS1 band at 2.9 kb is replaced by a 7.6 kb band in AM512, with the other band at 5.2 kb conserved in both strains. Using the NcoI(119)-NcoI(123) fragment as probe, strain AKS1 exhibits one band at 25.5 kb and one at 17.9 kb in the SphI and XhoI digests, 20 respectively, whereas, in addition to these bands, strain AM512 shows one SphI band at 5.5 kb and one XhoI band at 7.6 kb. In this way, it is established that the eryAKS1 allele has been converted into the eryAM512 allele in strain AM512.

25

### Example 44

### Isolation and purification of 14[1(1-hydroxypropyl)]erythromycin A

At harvest the pH of the fermentation of AM512 is adjusted to 9.5

and the mixture is extracted twice with equal volumes of ethylacetate.

The combined ethylacetate extracts are washed with water, dried and partitioned in a heptane, methanol, water (5:5:1, v/v/v) system. The lower (methanolic phase) is washed with an equal volume of heptane and is concentrated to a residue. This is chromatographed on a Sephadex LH-20 column in methanol and fractions containing the desired 14[1(1-hydroxypropyl)]erythromycin A are concentrated and further purified by chromatography on an Ito Coil Planet Centrifuge in a system consisting of n-heptane, benzene, acetone, isopropanol, 0.65 M, pH 7.0 aqueous

3

S

Ē

.

potassium phosphate buffer (5:10:2:3:5, v/v/v/v). Fractions containing the desired product are concentrated to a solid residue and partitioned between methylene chloride and dilute (pH 9.5) ammonium hydroxide. The organic layer is washed with water and concentrated to yield 14[1(1-hydroxypropyl)]erythromycin A as a white solid.

Although the present invention is described in the Examples listed above in terms of preferred embodiments, they are not to be regarded as limiting the scope of the invention. The above descriptions serve to illustrate the principles and methodologies involved in creating the three 10 types of mutations that can be introduced into the eryA segment of the Sac. erythraea chromosome that result in the synthesis of novel polyketide products. Although single Type I alterations, leading to the production of 5-oxo-5,6,-dideoxy-3α-mycorosyl erythronolide B, 11-oxo-11deoxyerythromycin A, 7-hydroxyerythromycin A, 7-oxo-15 7deoxyerythromycin A, 5-desosaminyl-3-oxo-3-deoxyerythronolide A, and Δ-6,7-anhydro-6-deoxyerythromycin A are specified herein, it is obvious that other Type I changes can be introduced into the eryA segment leading to novel polyketide structures. Among the additional Type I alterations that can be obtained are those in which two or more modules 20 are affected leading to the synthesis of novel polyketides. Examples of combinations of two Type I alterations leading to useful compounds include but are not limited to: mutants of the the  $\beta$ -ketoreductase of module 2 (KR2) and the  $\beta$ -ketoreductase of module 4 (KR4) leading to the formation of 7,11-dioxo-7,11-dideoxyerythromycin A; mutants of KR2 and 25 the  $\beta$ -ketoreductase of module 6 (KR6) leading to the formation of 3,11dioxo-3,11-dideoxy-5-desosaminylerythronolide A; mutants of KR2 and the dehydratase of module 4 (DH4) leading to the synthesis of 7-hydroxy-11-oxo-11-deoxyerythromycin A; mutants of KR2 and the enoylreductase 30 of module 4 (ER4) leading to the synthesis of  $\Delta$ -6,7-anhydro-11-oxo-11deoxyerythromycin A; mutants of KR4 and KR6 leading to the synthesis of 3,7-dioxo-3,7-dideoxy-5-desosaminylerythronolide A; mutants of KR6 and DH4 leading to the synthesis of 3-oxo-3-deoxy-5-desosaminyl-7hydroxyerythronolide A; mutants of KR6 and ER4 leading to the synthesis of 3-oxo-3-deoxy-5-desosaminyl- $\Delta$ -6,7-anhydroerythronolide A. Examples 35 of combinations of three Type I alterations leading to the synthesis of novel polyketides include but are not limited to: mutants of KR2, KR4 and KR6 leading to the synthesis of 3,7,11-trioxo-3,7,11-trideoxy-5-

غ

Ē

Ť

ŝ

5

desosaminylerythronolide A; mutants of KR2, KR6 and DH4 leading to the synthesis of 3,11-dioxo-3,11-dideoxy-5-desosaminyl-7-hydroxyerythronolide A; mutants of KR2, KR6 and ER4 leading to the synthesis of 3,11-dioxo-3,11-dideoxy-5-desosaminyl-D-6,7-anhydroerythronolide A. All combinations of two or three Type I mutants, the Sac. erythraea strains that carry said combinations and the corresponding polyketides produced from said strains, therefore, are

included within the scope of the present invention. Although the Type II mutants specified herein have been constructed in the  $\beta-ketoacyl$  ACP synthase of module 1 (KS1) and the  $\beta-$ 10 ketoacyl ACP synthase of module 2 (KS2), other Type II mutants can be constructed in other domains to result in the synthesis of novel polyketide structures upon feeding with appropriate substrate analogs. Other Type II mutants include but are not limited to: inactivation of the either of the acyltransferases or acyl carrier proteins of module 1, or the 15 acyltransferase or acyl carrier protein of module 2, the  $\beta$ -ketoacyl ACP synthase, acyltransferase or acyl carrier protein of module 3, module 4 or module 5. Furthermore, compounds other than (2S,3R,4S,5S)3,5dihydroxy-2,4-dimethylhexanoic acid-ethyl thioester and (2S,3S,4S,5S)2,4dimethyl-3-fluoro-5-hydroxyhexanoic acid-ethyl thioester specified herein 20 can be synthesized and fed to strains AKS1 or AKS2 specified herein or other strains that carry other Type II mutations to result in the creation of novel polyketides that are within the scope of the present invention.

Although two examples of Type III alterations are specified herein, it is apparent to those skilled in the art that many other examples of Type 25 III changes are possible. Strains of Sac. erythraea carrying changes of this type offer the very high potential for the production of novel polyketides of specified structure, since they do not require synthetic substrates as do Type II mutants and they are not limited to the formation of derivatives of erythromycin, as in the case of Type I mutants. In the embodiments of 30 Type III mutants specified herein, we have illustrated how a second copy of a complete module can be introduced at a desired position by gene conversion to result in the synthesis of 14-(1-propyl)erythromycin A or 14-[1(1-hydroxypropyl])erythromycin A. These alterations make use of the high conservation and simultaneous lack of specificity of the  $\beta$ -ketoacyl 35 ACP synthases of modules 1 and 2, thereby making possible the construction of hybrid  $\beta$ -ketoacyl ACP synthase functions consisting of portions of proteins derived from different modules. Those skilled in the

\*

ŝ

30

35

art understand, therefore, that it is possible, as exemplified for KS1 and KS2, to delete a small portion of the  $\beta$ -ketoacyl ACP synthase of other modules and to construct strains carrying such alterations which can then be employed as hosts for introducing at the deleted β-ketoacyl ACP synthase location a second copy of any homologous module. Furthermore, as exemplified herein, it is also possible to delete any segment of eryA by ligation of two non-contiguous PCR-generated fragments and to subsequently construct strains, therefore, devoid of any or all portions of any module. Such strains deleted of a full module can be employed for reintroduction of either the same or a different module at a 10 different location. It is possible, therefore, to determine the novel structures desired and then create a series of Sac. erythraea strains containing the corresponding arrangements of eryA modules that would produce said novel structures that are included within the scope of the present invention. Additional examples of novel compounds produced from the construction of Type III alterations include but are not limited to 11-deoxyerythromycin, resulting from the insertion of the eryA segment encoding DH4 and ER4 in module 2.

Moreover, it will also be apparent that two or more modules can be excised and introduced into various sites of the Sac. erythraea chromosome to produce novel polyketides of predicted structure such as the introduction of the eryA segment encoding DH4 and ER4 in both module 1 and module 2 to result in the production of 14(R)[1-hydroxypropyl]11-deoxyerythromycin A. All combinations, therefore, of Type III alterations and the strains of Sac. erythraea that carry said alterations as well as the polyketides produced from said strains are included within the scope of the present invention.

In addition, it is also possible to create combinations of Type I, Type II and Type III alterations and insert such alterations into Sac. erythraea to produce novel polyketides. Examples of such combinations include but are not limited to the following. The combination of a Type I alteration, such as an alteration in DH4 and a Type II alteration, such as a mutation in the KS1 to result in the formation of (14S,15S)14-[1-hydroxyethyl]-7-hydroxyerythromycin A when the strain of Sac. erythraea carrying such alterations is fed with the compound (2S,3R,4S,5S)3,5-dihydroxy-2,4-dimethylhexanoic acid ethyl ester. The combination of a Type I alteration, such as an alteration in DH4 and a Type III alteration, such as found in Sac. erythraea strain AM412, wherein a copy of the DNA segment of

Ē

Ź,

ĵ.

10

module 4 is introduced in module 1, such that the *Sac. erythraea* strain so constructed produces the compound 7-hydroxy-14-propylerythromycin A. All combinations of two or more alterations of Type I, Type II and Type III alterations, the *Sac. erythraea* strains that carry such alterations, and the polyketides produced from such strains are included within the scope of the present invention. It will also occur to those skilled in the art that novel structures can be produced by altering the specificity of the acyltransferase functions in any module. Examples include: replacement of the acyltransferase domains of modules 1, 2, 3, 4, 5, or 6 in *eryA* with those of modules 4, 4, 2, 2, 2, and 4, respectively, to result in the production of 12-epierythromycin A,10-epierythromycin A, 8-epierythromycin A, 6-epierythromycin A, 4-epierythromycin A and 2-epierythromycin A, respectively, that are included within the scope of the present invention.

It should be emphasized that the introduction of an entire eryA 15 module at a different location, as exemplified for the construction of Sac. erythraea strains AM412 and AM512 in Examples 29 and 35, respectively, does not rely on homologous recombination between the incoming eryA module and the host chromosome. Rather, gene conversion of the host allele with the eryA allele residing on the multicopy plasmid requires 20 DNA sequences homologous to the host allele flanking the incoming module. Thus, any module carrying the desired specificities, either from homologous or heterologous sources, can be employed in gene conversion of the host allele, provided that is flanked by segments of homology. It will occur to those skilled in the art, therefore, that, given 25 the large number of natural polyketide molecules existing, a wide variety of additional novel molecules of predicted structure can be produced in Type III mutants containing an additional module of desired specificities or where an endogenous module is replaced by an exogenous one. The length of the acyl chain can be easily controlled by suitably changing the 30 number of modules involved in its synthesis. Similarly, the introduction of keto, hydroxy, enoyl, or methylene groups at specific points along the acyl chain can be easily achieved by introducing the proper b-carbon processing functions (β-ketoreductase, dehydratase and enoylreductase) in the required modules. Exogenous modules constitute the source of 35 specificities for starter and extender units other than those employed by Sac. erythraea for erythromycin biosynthesis, making it thereby possible to employ, for example, malonylCoA or (2R)- or (2S)ethylmalonylCoA, etc.

•

ъ

Â

as extender units, and acetyl CoA, butyryl CoA, etc. as the starter unit. The result will be the formation of erythromycin analogs containing the desired functional groups and side chains with the desired stereochemistry. As an extension of the examples reported with eryA, the construction of a Sac. erythraea strain carrying a heterologous module inserted into eryA requires: (i) cloning of the genes from any other Actinomyces producing a polyketide with desired structural features; (ii) mapping of the modular organization of the cloned genes by low stringency hybridization and restriction analysis; (iii) locating the module carrying the desired specificities by partial sequencing; (iv) precise excision 10 of the desired genetic element and cloning into a vector suitable for gene conversion; (v) construction and transformation of a Sac. erythraea strain suitable for gene conversion and screening for the novel compound. Any module, or portion thereof, can thus be precisely excised from the genome of a polyketide-producing microorganism and introduced into suitable 15 Sac. erythraea strains to create a novel polyketide of predicted structure. Thus, replacement of the acyltransferase segments of modules 1, 2, 3, 4, 5, or 6 in eryA with the acyltransferase segment specific for malonyl CoA, such as can be found in the polyketide synthase genes for the synthesis of pikromycin in Streptomyces venezuelae, to result in the synthesis of 12-20 norerythromycin A, 10-norerythromycin A, 8-norerythromycin A, 6norerythromycin A, 4-norerythromycin A and 2-norerythromycin A, respectively, that are included within the scope of the present invention. In addition, replacement of the acyltransferase segments of modules 1, 2, 3, 4, 5, or 6 in eryA with an acyltransferase specific for (2R)-ethylmalonyl 25 CoA, such as can be found in the polyketide synthase genes for the synthesis of spiramycin in Streptomyces ambofasciens, will result in the formation of 12-homoerythromycin A, 10-homoerythromycin A, 8epihomoerythromycin A, 6-epihomoerythromycin A, 4epihomoerythromycin A and 2-homoerythromycin A, respectively, all of 30 which are included within the scope of the present invention. Similarly, introduction of acyltransferase segments carrying desired specificities for the starter or extender unit into eryA DNA that results in the synthesis of novel compounds are included within the scope of the present invention. The erythromycin analogs produced by the method of this invention are 35 structurally similar to known antibacterial and prokinetic agents.

It will also occur to those skilled in the art that genetic manipulations described herein need not be limited to Sac. erythraea.

\$

Ŧ

Ŝ

Suitable hosts are any other polyketide-producing Actinomyces where DNA can be precisely inserted into the chromosome. Hence, the choice of a convenient host is based solely on the relatedness of the novel polyketide to a natural counterpart so as to minimize the number of module rearrangements required for its biosynthesis. Therefore, Type I, Type II and Type III alterations can be constructed in other Actinomyces employing either endogenous or exogenous modules to produce novel polyketides employing strategies analogous to those described herein for Sac. erythraea. Thus all Type I, Type II or Type III mutations or various combinations thereof constructed in other actinomycetes according to the 10 principles described herein, and the respective polyketides produced from such strains, are included within the scope of the present invention. Examples of polyketides that can be altered by creating Type I, Type II or Type III changes in the producing microorganisms include, but are not limited to macrolide antibiotics such as erythromycin, tylosin, spiramycin, 15 etc.; ansamacrolides such as rifamycins, maytansines, etc.; polyketide antibiotics such as tetracycline; polyethers such as monesin, salinomycin, etc.; polyenes such as candicidin, amphothericins; immunosuppressants such as FK506, ascomycin, rapamycin, etc. and other complex polyketides such as avermectin. 20

Whereas the novel derivatives or modifications of erythromycin described herein have been specified as the A derivatives, such as 7hydroxyerythromycin A, 11-oxo-11-deoxyerythromycin A, 14[1(1hydroxypropyl)]erythromycin A, etc., those skilled in the art understand that the wild type strain of Sac. erythraea produces a family of 25 erythromycin compounds, including erythromycin A, erythromycin B, erythromycin C and erythromycin D. Thus, modified strains of Sac. erythraea, such as strain AKR2, for example, would be expected to produce the corresponding members of the 11-oxo-11-deoxyerythromycin family, including 11-oxo-11-deoxyerythromycin A, 11-oxo-11-deoxyerythromycin 30 B, 11-oxo-11-deoxyerythromycin C, and 11-oxo-11-deoxyerythromycin D. Similarly, strain AM412 would be expected to produce not only 14(1propyl)erythromycin A but also the other members of the 14(1propyl)erythromycin family including 14(1-propyl)erythromycin B, 14(1propyl)erythromycin C and 14(1-propyl)erythromycin D. Similarly, all 35 other modified strains of Sac. erythraea described herein that produce novel erythromycin derivatives would be expected to produce the A, B, C, and D forms of said derivatives. Therefore, all members of the family of

<u> 3</u> -

Ŕ

F

each of the novel polyketides described herein are included within the scope of the present invention.

Variations and modifications of the methods for obtaining the desired plasmids, hosts for cloning and choices of vectors and segments of eryA DNA to clone and modify, other than those described herein that 5 result in substantially the same strains and same products as those described herein will occur to those skilled in the art. For example, although we have described the use of the plasmids pWH3 and pWHM4 as E. coli-Sac. erythraea shuttle vectors, other vectors can be employed wherein all or part of pWHM3 or pWHM4 is replaced by other DNA segments that function in a similar manner, such as replacing the pUC19 component of pWHM3 and pWHM4 with pBR322, available from BRL, employing different segments of the pIJ101 or pJV1 replicons in pWHM3 and pWHM4, respectively, or employing selectable markers other than thiostrepton- and ampicillin-resistance. These are just few of a long list of 15 possible examples all of which are included within the scope of the present invention. Similarly, the segments of the eryA locus subcloned into pWHM3 for generating strains AKS1, AKS2, etc. specified herein can readily be substituted for other segments of different length encoding the same functions, either produced by PCR-amplification of genomic DNA or 20 of an isolated clone, or by isolating suitable restriction fragments from Sac. erythraea. In the same way, it is possible to create eryA strains carrying mutations functionally equivalent to those described herein by deleting different portions of the corresponding genes, by creating insertions into them, or by site-directed mutagenesis of specific nucleotide residues. 25 Moreover, Sac. erythraea strains with mutant alleles other than the  $\beta$ ketoacyl ACP synthase portions of eryA can be employed as hosts for gene conversion; Type III mutants can be constructed by double reciprocal crossover as exemplified for Type I and Type II mutants rather than by the gene conversion method described herein. Additional modifications 30 include changes in the restriction sites used for cloning or in the general methodologies described above. All such changes are included in the scope of the invention. It will also occur to those skilled in the art that different methods are available to ferment Sac. erythraea, to extract the novel polyketides specified herein, and to synthesize substrate analogs, 35 and that all such methods are also included within the scope of the present invention.

It will be apparent that many modifications and variations of the invention as set forth herein are possible without departing from the spirit and scope thereof, and that, accordingly, such limitations are imposed only as indicated by the appended claims.

5

3

ō

۶,

30

ક

## What is claimed is:

- 1. A method for directing the biosynthesis of specific polyketide analogs by genetic manipulation of a polyketide-producing microorganism, said method comprising the steps of:
- (1) isolating a polyketide biosynthetic gene-containing DNA sequence;
- (2) identifying enzymatic activities associated within said genecontaining DNA sequence;
- (3) introducing one or more specified changes into said genecontaining DNA sequence which codes for one of said enzymatic activities resulting in an altered DNA sequence;
  - (4) introducing said altered DNA sequence into a polyketideproducing microorganism to replace the original sequence;
- (5) growing a culture of the altered microorganism under conditions suitable for the formation of the specific polyketide analog; and (6) isolating said specific polyketide analog from the culture.
- 2. The method of claim 1 wherein said polyketide biosynthetic genecontaining DNA sequence comprises genes which encode the enzymatic activities comprising a polyketide synthase.
- 3. The method of claim 2, wherein said polyketide synthase enzymatic activities comprise β-ketoreductase, dehydratase, acyl carrier protein, enoylreductase, β-ketoacyl ACP synthase, and acyltransferase.
  - 4. The method of claim 1 wherein said alteration which occurs in the DNA sequence results in the inactivation of one or more enzymatic activities involved in the processing of the  $\beta$ -carbonyl of said polyketide.
  - 5. The method of claim 4, wherein said inactivated enzymatic activities affecting the processing of the  $\beta$ -carbonyl of said polyketide comprise  $\beta$ -ketoreductase, dehydratase, and enoylreductase.
- 35 6. The method of claim 4 wherein said alteration in the DNA sequence results in the addition of one or more enzymatic activities involved in the  $\beta$ -carbonyl processing of said polyketide.

WO 93/13663 PCT/US92/00427

40

7. The method of claim 6 wherein said additional enzymatic activities are selected from the group consisting of  $\beta$ -ketoreductase,  $\beta$ -ketoreductase and dehydratase, and  $\beta$ -ketoreductase, dehydratase and enoylreductase.

3

Ê

Ĭ.

- 5 8. The method of claim 1 wherein said alteration occurring in the DNA segment results in the inactivation of one or more enzymatic activities involved in the condensation of carbon units to the nascent polyketide structure.
- 10 9. The method of claim 8 wherein said enzymatic activities affecting the condensation of carbon units to the nascent polyketide structure comprise  $\beta$ -ketoacyl ACP synthase, acyl carrier protein and acyltransferase.
- 10. The method of claim 1 wherein said alteration in the DNA15 sequence results in the change of the length of the polyketide synthesized.
  - 11. The method of claim 10 wherein said alteration results in the increase of the length of the polyketide.
- 20 12. The method of claim 11 wherein said alteration comprises the addition of DNA sequences encoding the enzymatic activities consisting of acyltransferase, acyl carrier protein and β-ketoacyl ACP synthase.
- 13. The method of claim 10 wherein said alteration results in the decrease of the length of the polyketide.

30

- 14. The method of claim 13 wherein said alteration consists of the deletion of a DNA segment between two sequences encoding corresponding enzymatic activities.
- 15. The method of claim 14 wherein said corresponding enzymatic activities are selected from the group consisting of  $\beta$ -ketoreductases, dehydratases, acyl carrier proteins, enoylreductases,  $\beta$ -ketoacyl ACP synthases, and acyltransferases.
- 16. The method of claim 1 wherein said alteration consists in the replacement of the DNA segment encoding an acyltransferase with a DNA segment encoding an acyltransferase of different specificity.

.

\$

'n

- 17. The method of claim 1 wherein said DNA sequence is isolated from a species from the Actinomycetales family.
- 5 18. The method of claim 17 wherein said DNA sequence is isolated from a genus selected from the group consisting of Actinomyces, Dactylosporangium, Micromonospora, Nocardia, Sac., Streptoverticillium, and Streptomyces.
- 10 19. The method of claim 17 wherein said genus is selected from the group consisting of Saccharapolyspora and Streptomyces.
  - 20. The method of claim 19 wherein said genus is Saccharapolyspora and the species is erythraea.
- 15
  21. The method of claim 19 wherein said genus is Streptomyces and the species is hygroscopicus.
- The method of claim 1 wherein said polyketide is selected from the group consisting of macrolides, tetracyclines, polyethers, polyenes, ansaymcins and derivatives or analogs thereof.
  - 23. The method of claim 22 wherein said polyketide is a macrolide.
- 25 24. The method of claim 23 wherein said macrolide is an erythromycin.
  - 25. The method of claim 24 wherein said erythromycin analog is selected from the group consisting of 11-oxo-11-deoxyerythromycin A, 7-hydroxyerythromycin A, 6-deoxy-7-hydroxyerythromycin A, 7-
- 30 oxoerythromycin A, 3-oxo-3-deoxy-5-desosaminylerythronolide A, Δ-6,7-anhydroerythromycin A, ((14S,15S)14(1-hydroxyethyl)erythromycin A, 11-epifluoro-15-norerythromycin A, 14-(1-propyl)erythromycin A, 14(1-propyl)erythromycin A, and 14[1(1-hydroxypropyl)]erythromycin A.
- 3 5 26. The method of claim 1 wherein said DNA sequence, designated eryA, encodes the enzymatic activities associated with the formation of 6-deoxyerythronolide B.

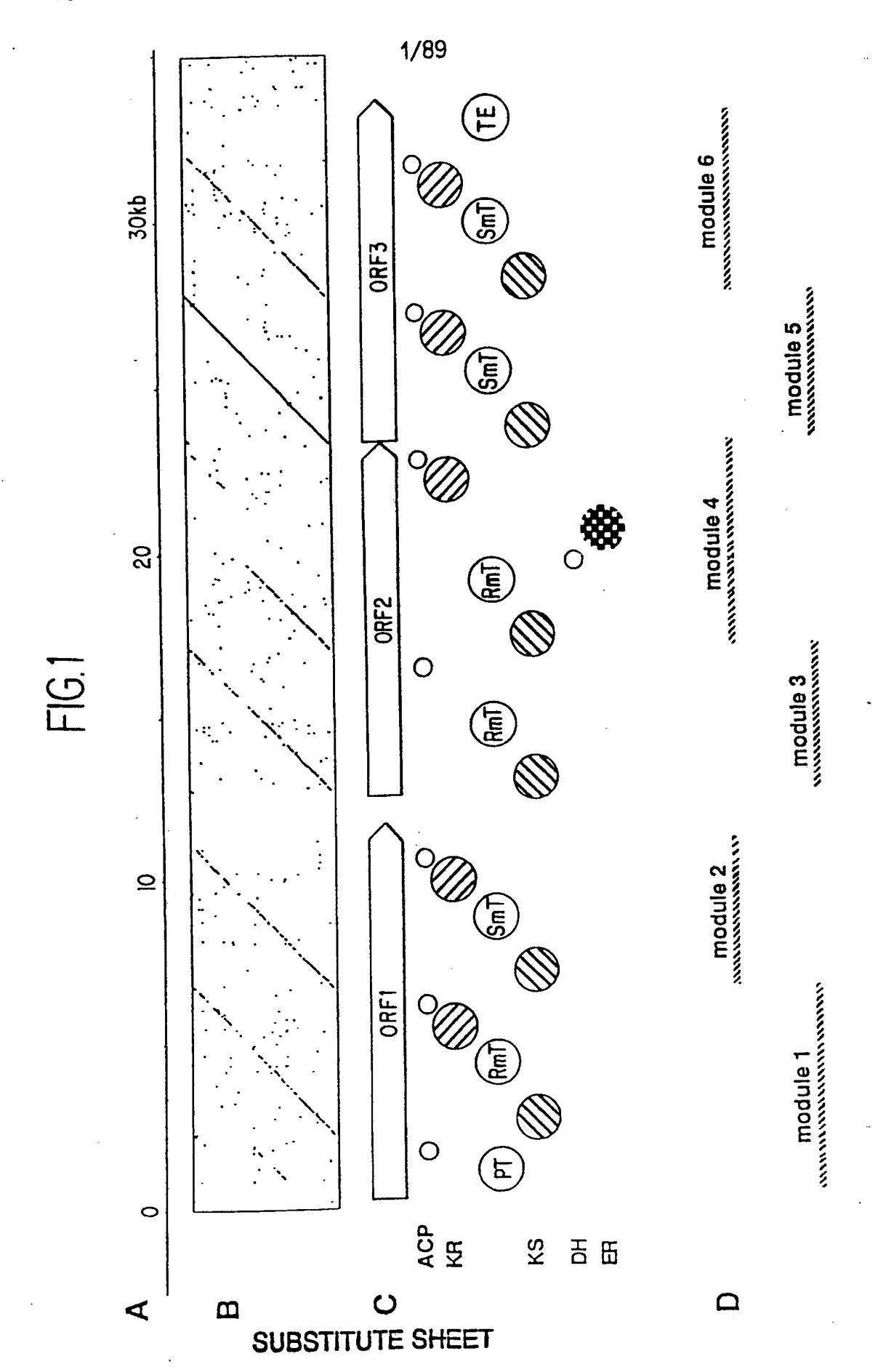
WO 93/13663 PCT/US92/00427

42

Ž

Â,

- 27. The method of claim 26 wherein said DNA sequence comprises: the DNA sequence of Figure 2.
- 28. The method of claim 1 wherein said gene-containing DNA sequence encodes one or more enzymatic activities in the rapamycin biosynthetic pathway.
  - 29. The method of claim 23 wherein said macrolide is a rapamycin analog.
- 30. A compound selected from the group consisting of 7-hydroxyerythromycin A; 6-deoxy-7-hydroxyerythromycin A; 7-oxoerythromycin A, 3-oxo-3-deoxy-5-desosaminyl-erythronolide A; Δ-6,7-anhydroerythromycin A; ((14S,15S)14(1-hydroxyethyl)erythromycin A; 11-epifluoro-15-norerythromycin A; 14-(1-propyl)erythromycin A; 14(1-propyl)erythromycin A; and 14[1(1-hydroxypropyl)]erythromycin A.



ż

**÷** 

ð

ź

	- -					3.2					
	0		0		0	FI(	09		20		0 8
ı	108	1	114	ı	120	I	12(	i	133	ì	H 3
囧	00 + -	>	H +	ഗ	) † †	缸	3CG + - +	凹	AGA	Н	CGC
Ĺτι	500	æ	CTGGCG	æ	090	K	) ) ) ) )	R R	CGACGAGA	回	GTC
æ	GCG 	民	CHG	3	CGC	K	SGAC	S	Öi	Q	ACGGTC
DC.	GTC.	S	GCGCT	ы	+: +:	A	TGTGGAG	3	CCGC	æ	HCA ++
· 4	AGCAGTCGCGCCG	a	292	æ	GCTGG(	1	CGC1	H	HCH(	လ	3998
ບ	CGA	回	255	æ	CGA	回	2002	A	CO I	H	)   
4	H C I	Ω	GCT	Н	990,	Ö	7000	<b>4</b>	rage	K	FGG + 1
Ω	0HU +	Ø	GTC +	ഗ	CAT-	Н	3000	K	Ö	>	IGC.
E	GGA	Ω	GAC	ĘH	CAG	ഗ		K.	CAG(	. <b>Æ</b>	TGG.
A	CCT	Н	GCA	ø	CCA	##	3000	K	TGG(	K	ACG
.eq	-+-	>	-+-	>	 -+	U	CGACG(	K	ACAT	Σ	ACG +
A	GCA	a	200	K	1661	>	( ) 1	Ω	900	Ω	ACG
Įz.	000 1	, K	GTT	मि	3CGT	>	3CG(	Æ	ACG	O	0 0 0 0
>	1901 +	ы	;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;	Ч	ACGC	A	GTG(	Æ	GCA ++	Z	GGT -+-
DC,	GAC	E	2000	K	CCG	Ω	) CCC	Ŋ	7GG	U	0
S	CTC	3	VGCC	Си	CC	Д	CGG	K	TGG 	>	T C C
回	(D)	Ω	GTCCA	Ø	rga(	H	GGTG( -+	æ	CCGT	Н	GCA +
<b>U</b>	GAC.	Ę	1001 1	>	SCGTG	>.	909	Ŋ	TTC	Δι	292
H	CG1	>	AGG7	>	ICG	Ö	TGT	U	TGA	<b>H</b> ¨	AGC 
H	AGCCCGTGACCGACTGG	Д	TCGAGGTCGTCCAGCCC	田	CCTTCGGCGTGACCCCC	मि	ACGTGTGCGGTGCGGCC	>	AGATGATTCCGTTGGTGG	Σ	TCGAGCCGCGCATCGCC
	1021		1081		1141		1201	•	1261		1321

ક

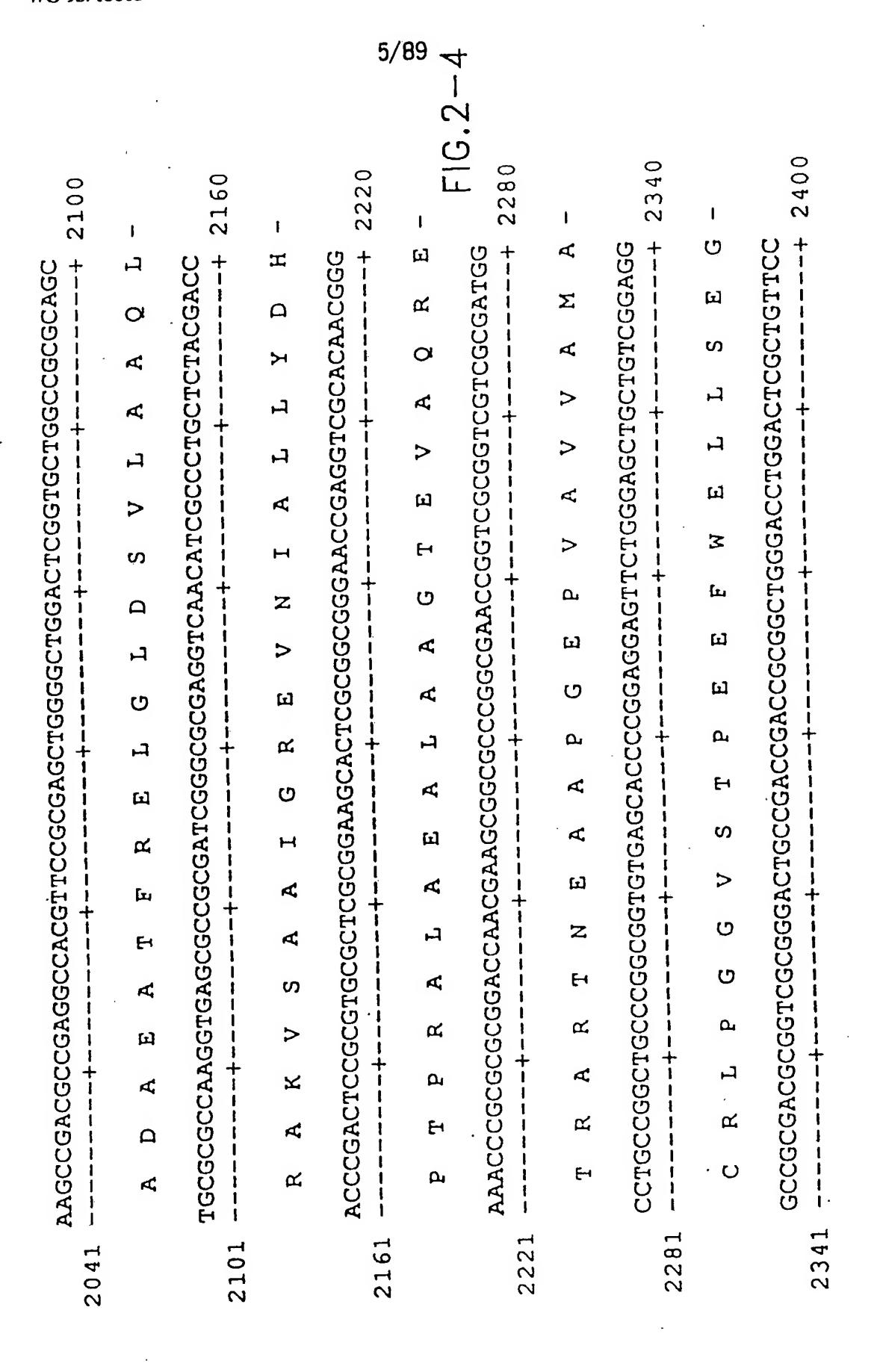
3/89  $\infty$ 回 ACTGGCGCCGCAGCTTCCGGCTGCGCTTCGACGAGGCGATCCGGTCCGCCTGG CCTGGCGTGGTTCGCCCCGGTGGCTCGGAGG S TCGG Z 只 GGCCGA TGTGTCGATGGCGCGCACTCGGCGCAGGT J. 回 S ACCGGTCGCGCCGGGTCCAGGAGCT A, S Ø U TGCGGTCGACACGCGGGAGCTGGT S U K 回 Z 民 Н C S Q 回 H  $\Omega_{1}$ 工 C A, 民 K A 只 W H Fi K K 0 3 Σ K L A, S 民 K Д ACGACATCGCCGAGGGGATGCGCTCGGC TGCCCTTCTACGCCAGCCTCACCGGAGG GCTCGGTTCTGCTGACCGGGTCGCCGGA CCGAGGGGTCCGCGCACAGGTCATCAA Ö K Z 闰 Ω  $\mathcal{O}$  $\alpha$ S H  $\boldsymbol{\sigma}$ ᆈ H  $\alpha$ > S 3 M, Ц  $\Sigma$ Q G K ī S G K H K S K, 回 K H ø, ≯ B > Н K ar. H G 3 Ω  $\Omega_{\mathbf{i}}$ 回 S 回 1621 1561 1501 1441 1381

?

1681	AGGTCGGTCCCGGCACGTCGAAGCGAGCCCGCACCCGGTGCTGGCCGCCGCGCTCC +++++++
	VGPGTFVEASPHPVLAALO-
1741	AGCAGACGCTCGACGCCGCTCCTCGGCCGCGGTGGTCCCGACGCTGCAACGCGGGC
	QTLDAEGSSAAVVPTLQRGQ-
1801	AGGCGGCATGCGGCTTCCTGCTGGCCGCGCCCAGGCGTTCACCGGCGGCGTGGCCG+++++++1860
	GGMRRFLLAAAQAFTGGVAV
1861	TCGACTGGACCGCCGACGACGTGGGGCCGAACCCGGCTCTCTGCCGGAGTTCGC
	DWTAAYDDVGPNPALCRSSR-
1921	GCCGGCCGAGGAGACGAGCCGGCCGAGTCCGGCGTCGACTGGAACGCGCCACCGCA 
	RPRRTSPRASTGTRHRT-
1981	CGTGCTGCGAGCGGCTCGTCGTCAACGGCGAGACCGCCGCGTTGGCGGGCCGCG
	C C E R L L A V N G E T A A L A G R E -

ŝ

Ź:



F

ġ.

Ā

```
6/89 <sub>LO</sub>
                                                                                                                                                               9
                                                                                                                                  00
                                                                                                     40
                                                                         80
                                            0
               60
                                                                                                                                  27
                                                                                                     9
                                                                         S
                                            S
               24
                                                                                                     2
                                                                         2
                                            2
                                                                                                                                                         GGAGGCCCGGCGATCAGCG
                                                                                                TACG
                                                                                                                                              H
                                                                                                                 G
                                                                                                                            TGATGACCGGTACGACCA
                                                                                    4
                                                                   CCTGGGAAGTGCTGGAACGGGGGGGGAATCCCGC
                            K
                                       工
           CCTGACCGAGG
                                                                                                                                              H
                                                                                                                 ×
                                                                                     Д
                            回
بنا
                                                                                                CGACCTCGTTGCAGGCCTCGCCCACTGGCGTGTTCGTCGGCCTGATCCCGCAGGAG
                                                                                                                                              H
                                                                                                                 回
                                                                                     H
                            H
H
                                                                                                                                              C
                                                                                                                 Q
                                                                                     \Omega
                                                         K
S
                                                                                                                                              H
                                                                                                                 Д
                                                                                     K
                            [LI
Ω
           GCACCAGCGCGGCGGG
                                                                                                                                              \Sigma
                                                                                                                 H
                                                                                     K
                                                         A,
                            U
                                                                                                                                              H
                                                                                                                             GCCCGCGGCTGGCCGAGGCGCGAAGGCGTCGAGGGCTACC
                                                                                     回
                                                         回
                            U
×
                                                                                                                  C
                                                                                     H
                                                         K
                            C
3
                                                                                                                                              U
                                                         PH
                             K
9
                                                                                                                                               回
                                                                                                                                                          ACACGC
                                                                                                                  FI
                                                                                      回
                                                         S
                             O
 民
                                                                                      3
                                                          \mathbf{X}
                             工
                                                                                                                                               C
            ACCCCGACCCCACGCGCTCGGGCACCGC
                                                                                                                  G
                                                                                                                                                           CGAGCGTCGCCTCCGGCCGCATCGCCT
                                                                     CGCAGCAGCGCTCATGCTCGAGCTCTC
                                                                                      S
                                                         5
                             K
 H
                                                                                                                                               回
                                                                                                                   H
                                                                                      H
                                                          [FI
                             H
                                                                                                                                                C
                                                                                                                   Д
                                                                                      回
                                                          [i4
                             C
 山
                                                                                                                                                \mathcal{O}
                                                                                                                   S
                                                                                      H
                                                          A
                              S
 G
                                                                                                                                                团
                                                                                                                   K
                                                                                      Σ
                                                          D<sub>4</sub>
                              K
 K
                                                                                                                                                K
                                                                                      Н
                                                                                                                   Q
                                                          Q
                              H
                                                                                                                                                \boldsymbol{\vdash}
                                                                                                                   Н
                                                                                      R,
                                                          [tı
                              р
                                                                                                                                                 \alpha
                                                                                      Ø
                                                                                                                   S
  Ω
                                                                                                                                                 Д
                                                                                                                   H
                                                                                      Q
                                                           H
  A,
                              A
                                                                                                                                                                  2701
                                                                                                                                     2641
                                                                                                        2581
                                                                           2521
                                               2461
                  2401
```

¥

¥

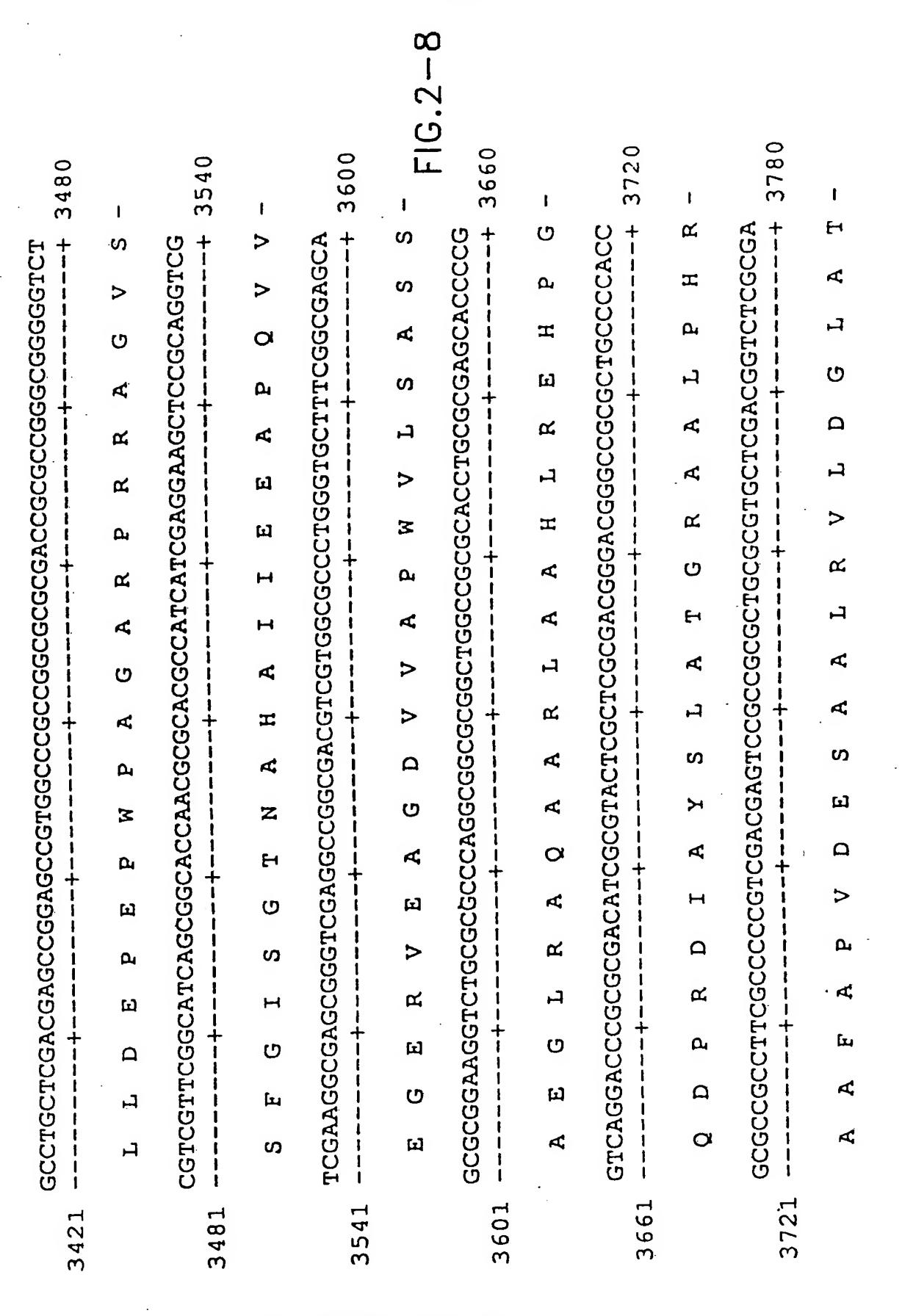
G

7/89 W 60 000 80 0 **C**1 30  $\infty$ 28 I 2 2 S CGGACGCCCCCCACCCCGGTGCTCGCCGTGCTCAGGGGGACGGCGGTCAACT A. S S H K GCGGCGAGTCGTCGCTGGCGATGGCGGCGGTGTCACGGTGATGCCGACGCCCGGCATGC CCGGC ÜL S  $\Sigma$ CL, S Ū Ū GGAGCG Ē TGGACACCGCGTGCTCGTCGCTGGTCGCGGTGCACCTGGCGTGCCAGTCGCT [L U TGGTGGACTTCAGCCGGATGAACTCGCTGGCGCCGGGACGGCCGGTGCAAGGC A. 回 K  $\Omega_{\mathbf{i}}$ S H × H Q Д Ū U O  $\Omega_{\mathbf{i}}$ O U K, K AGGCCCCGGGATGC  $\Sigma$ K 回 Ц  $\Sigma$ C Н H C Ω H I Ü K Z, Д G G A,  $\mathcal{O}$ H CCGGCGCCAACGGTTTCGGCATGGCCG  $\Box$ > ×  $\boldsymbol{\alpha}$ A S K Ы K 1 + + + - - -エ Σ Z Σ  $\dot{\mathbf{H}}$ S U  $\mathcal{O}$ Σ K S  $\alpha$ ------Z [T] 只 H S G K G S S  $\mathcal{O}$ S 1+1 民 Z بتا S K K A. K  $\Box$ 回 E > Ω U 9 Q S 3001 2941 2881 2821 2761

8/89 3300 60 3180 3120 34 3 32 3 I CCCGCACTCTGCACGCATCGGTCGAAGGAGATCGACTGGTCATCCGGTGCGATCA S K 民 GCGACCGCGAGCAGCTGCACCTGGGCTCGGTCAAGTCCAACCTCGGCCACACCCAGG AGATGGTGCTGGCGATGCGCGCGGGCACCCTTC Д G GCACCGGTACCCGACTCGGCCGATCGAGGCGCGGGCGCTGTTCGAGGCGTACGGGC Q AGCAGGCGCTGGCAGAGTCCGGGTCTCGGGCCCGCCGACATCGACGCCGTCGAGGCGCACG H Q U 工 H K H 7 K > 9 C K 工 pc, 딥 S C 回 > > S K H [14 K Q 3 Σ H Z K S K CL, H H X K U D K K Z ×  $\Sigma$ S 回 4 PI ഗ CGGCCGCCGGTGTTGCCGGCGTGATCA G H or, H ρι S 더 工  $\mathcal{O}$ Q  $\mathfrak{O}$ S 9 U S K Щ 回 Z 工 Q  $\alpha$ K S Н 回 H H  $\mathfrak{O}$ K H Д 3301 3361 3181 3121 3061

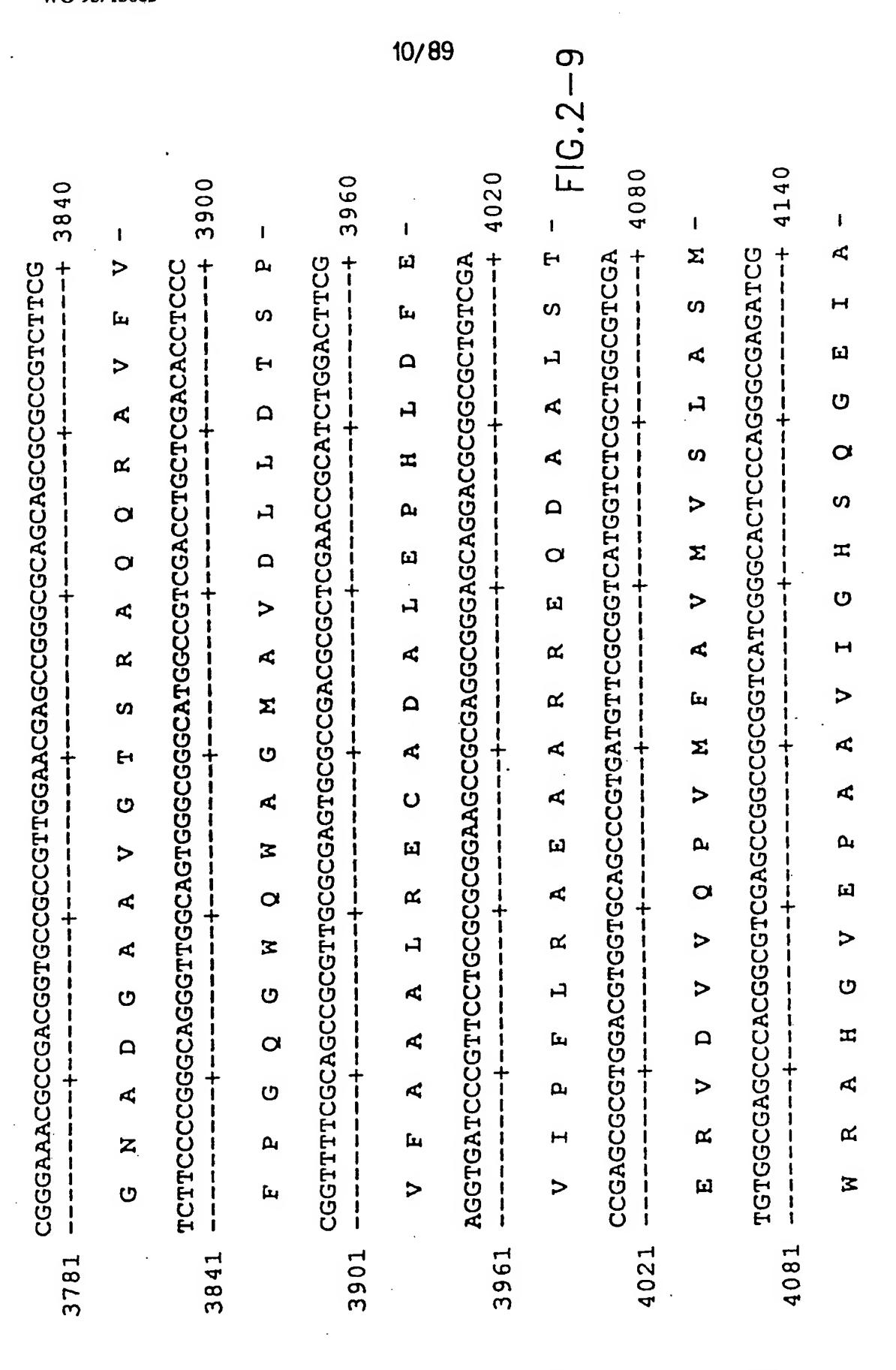
SUBSTITUTE SHEET

÷

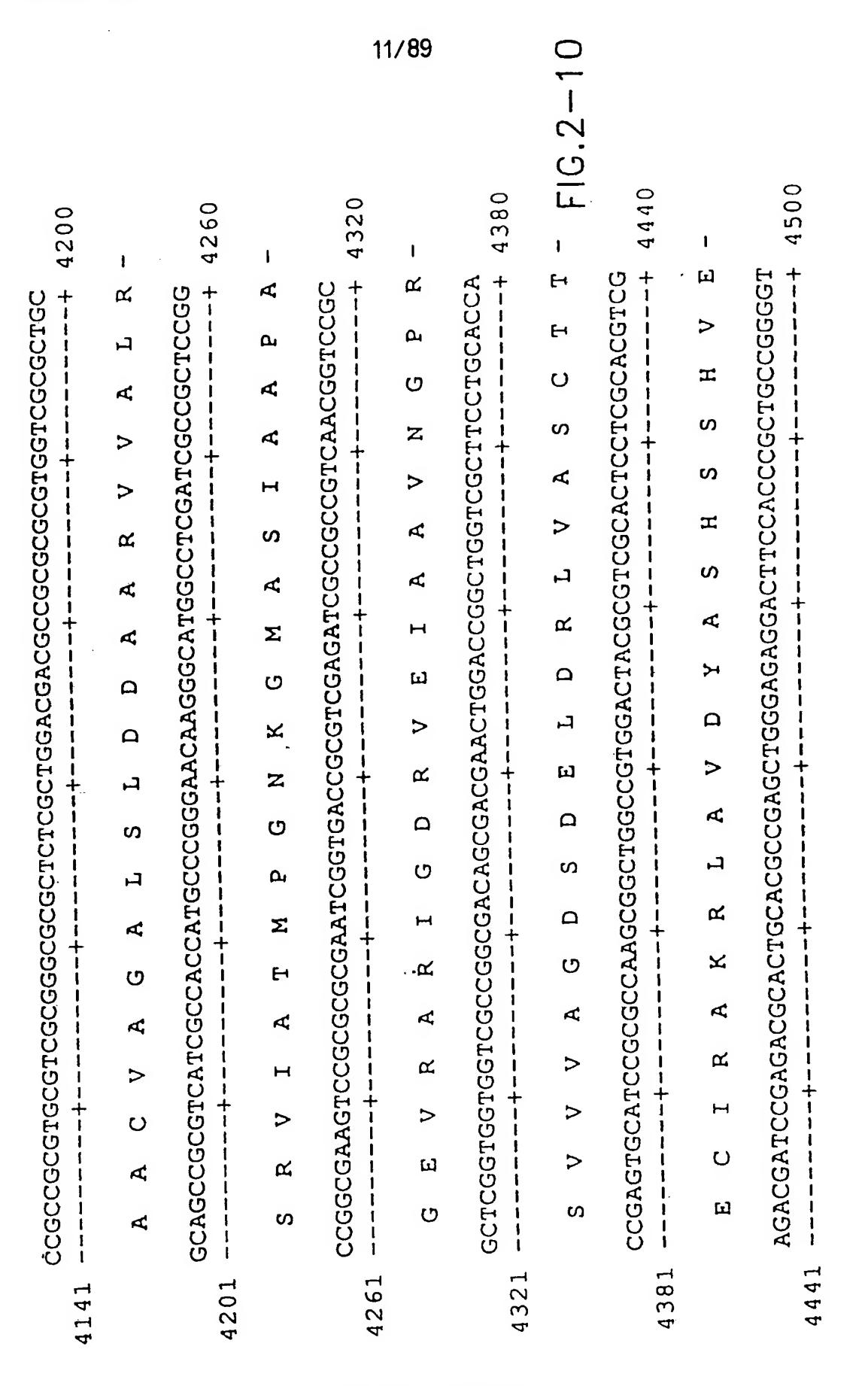


4

á.



**\*** 



S.

÷

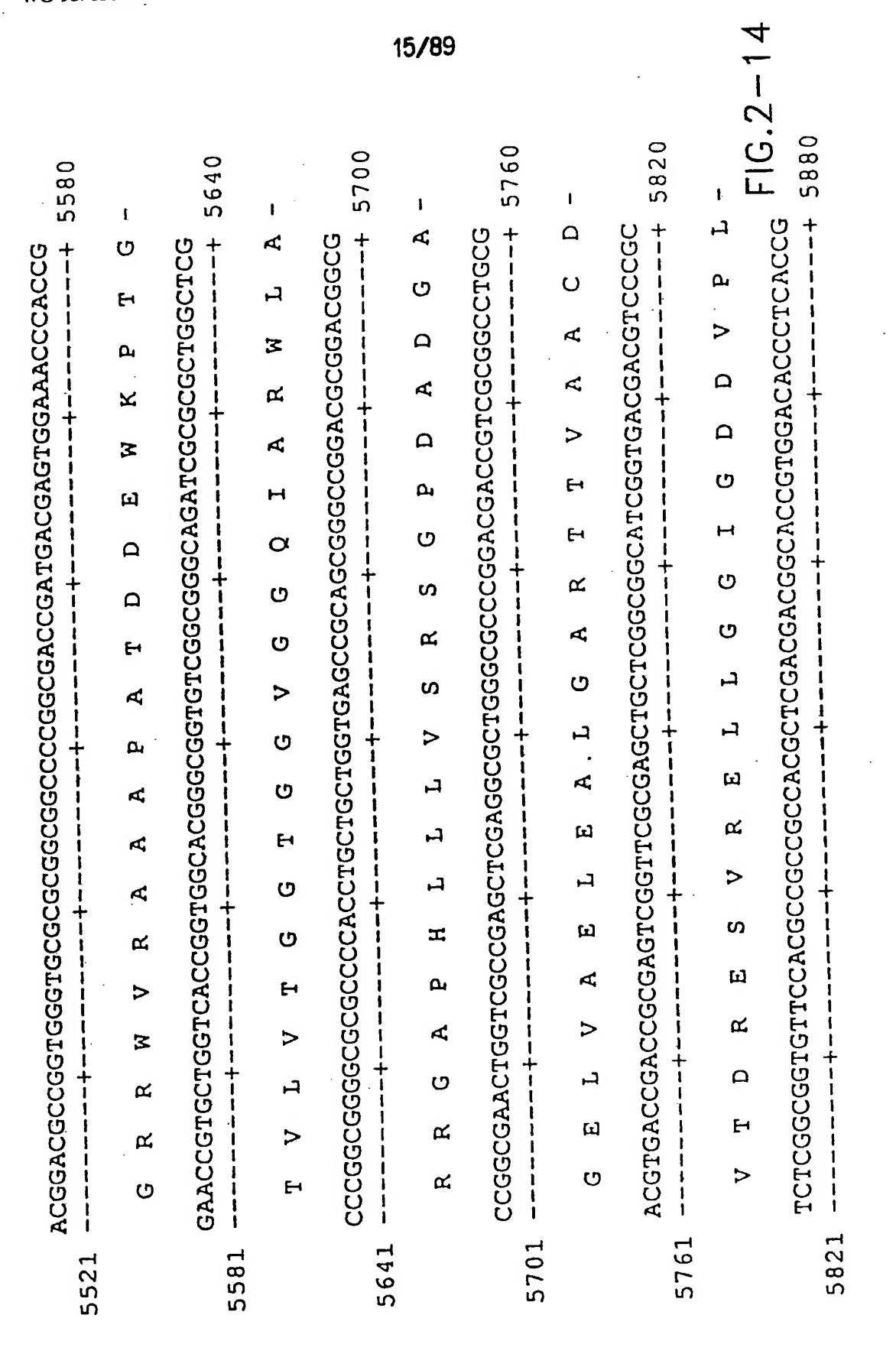
Ä

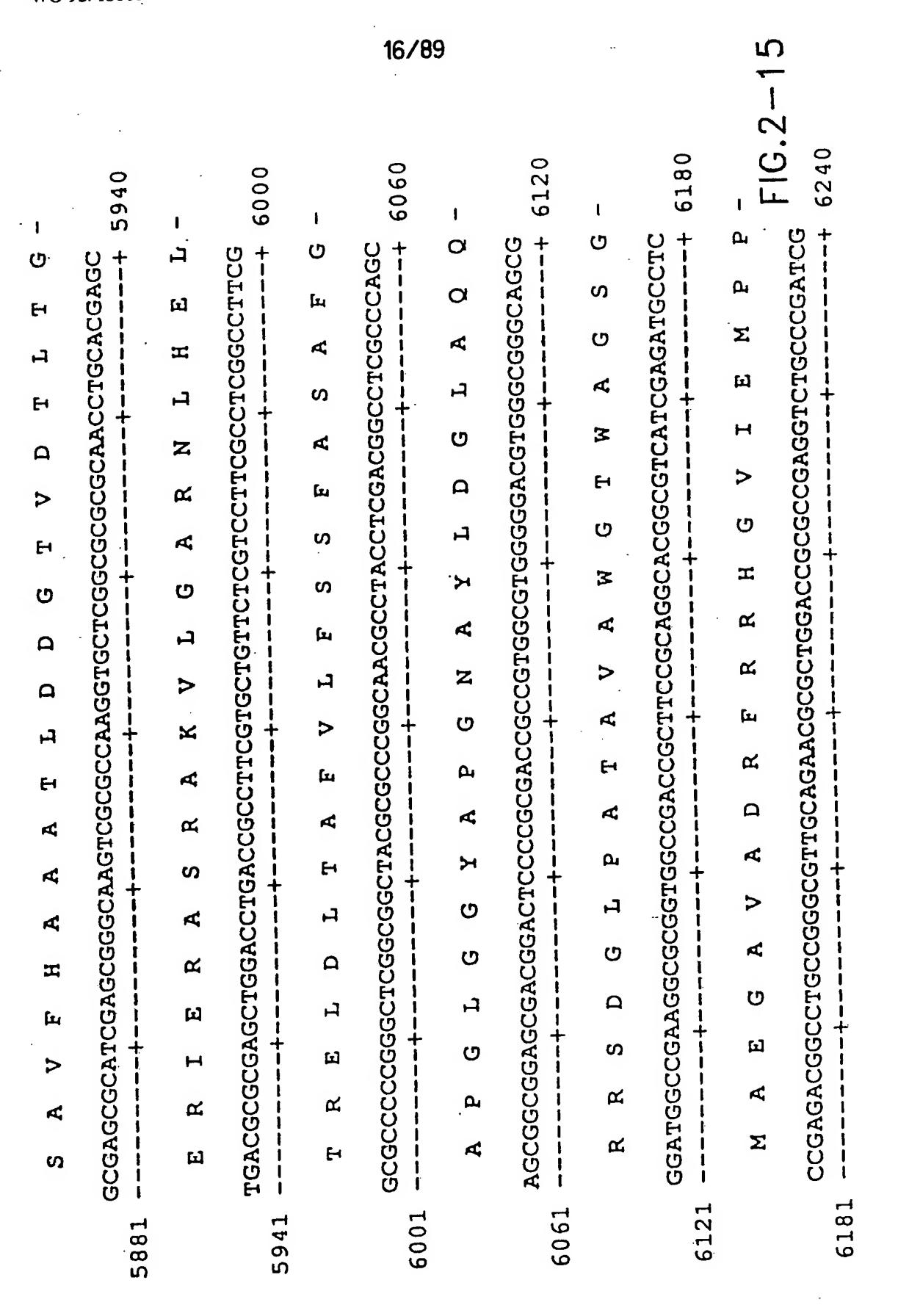
12/89 4800 680 20 4560 47 4 4 K CGGTGGACTGGGTGCACCTGGGCACCGGAGCACGCGGGGTGCCCTTGCCCACCT AAGCGCTCTCCCGCGCGTTCGCCGCCGGTGTCG C CCGACCTCTCCGCCATTCGCTGCGCGCG H AGGTCAGCGCACCCGATCCTCACCGCGCGA A 5 + H GCGCTGGACGCAGCCGGACGAGCTCGACGCCG > K K K G U 只 Ω щ K H H K H H K S 回 PI H 耳 H K D I A. H 24 D Д <u>Fu</u> K 工 K Q D S S K [t4 H 团 S K Z U K R 回 TCGTGCCCTTCTTCTCCACCGTCACCGG GTGACGGCAGCCTCGCGGACTTCGGCG TCGAGGAGATCGGCGACGGATCGGGCG 回 CCGAGCAGGGATATCGCACGTTCCTGG U 回 U G H K H K S [L [4 K 工  $\mathbf{Q}$ H H H K CL, Z S A. U × D, [LI S HC × K G 回 Q Z H 回 回 × H 4801 4621 4561 4501 468

					13/89			7	7	
I	4920	1	4980	I	5040	1	5100	ָ ( (	FIG. Z-	
<b>&gt;</b> 4	ტ + ე	山	H + +	ഥ	4 + C	H	00+	>.	9 9 1	K
E4 .		H		<b>o</b>	CGA	问	(A)	>	GGT	>
ρı	DDE.	S	ŬΙ	A,	GGACGAGA	Ω	ACT I	H	CGAGGT	ធ
H	000	Д	.GG1	Ö		K	CGAACTGGT	ជា	CGG(	r
<u>α</u>	SCGCCGGTCCAC	æ	CACCGGTG	EH	AACCGC	H	500	æ	CGGT(	>
>	) 1 1 1 1	K	3CC(	ርፈ	CGGA	ഗ	[-1	>	H	တ
K.	CGAACCGAAGCCGGTGGCG	>	CTACCGCATCGAGTGGCGGC	R.		K	+	民	GCTTCG +	æ
æ	3000 + - +	Д	37G( +	3	A H H	₩	GGC +	Æ		H
Æ	BAAG	×	GA(	ធ	GAA	×		ပ	909	α.
U	2000	Ċι	CAT(	н	GGCGAAGTACGC	K		æ	CGCGGAGCG	<u>屆</u>
EH	3GA2	<b>四</b>	3CG(	æ	GGT -+-	>	GTC -+-	တ		Æ
ט		H	CTA(	×	[H ]	H	CTGGAGTC	凹	CGAACT	H
H	CTGGCT	3	C	<b>K</b>	OFG I	Z	O I	H	CGA	ា
Ħ	CGT(	>	GCT(	H	CAC(+++++++++++++++++++++++++++++++++++	E	1 - + 1 - +	Æ	CGA-+	Ω
>	) 2 2 3 4 1	R.		A		<sub>ග</sub>	-4	团	E C C	ď
ഗ	GAC	回	FTC(	တ	CGA(	Ω	HCG I	CC.	990	Ŋ
凹	3000	K	SGT	>	SCH	ᆸ	000 1 1 1	æ	CTG	Ö
3	CAC	Ø	GAG	回	CCG(	K	) + () () 1	K	CCGC	CL.
Ω	TIC	Ĺij	GAC	Ω	) 2 2 2 2	K	CAC	₽	201	æ
>	ACCCGTTCCAGCGCGAGCG	ርፈ	AGGTCGACGAGGTTTCCGC	>	AACCCGCCCGGCTCGACGG	Д	CGAGCACCGCGGCTCGGGA	ဟ	TGGACGCCCGCTGCGGTCG	Ω
	4861		4921		4981		5041	_	5101	

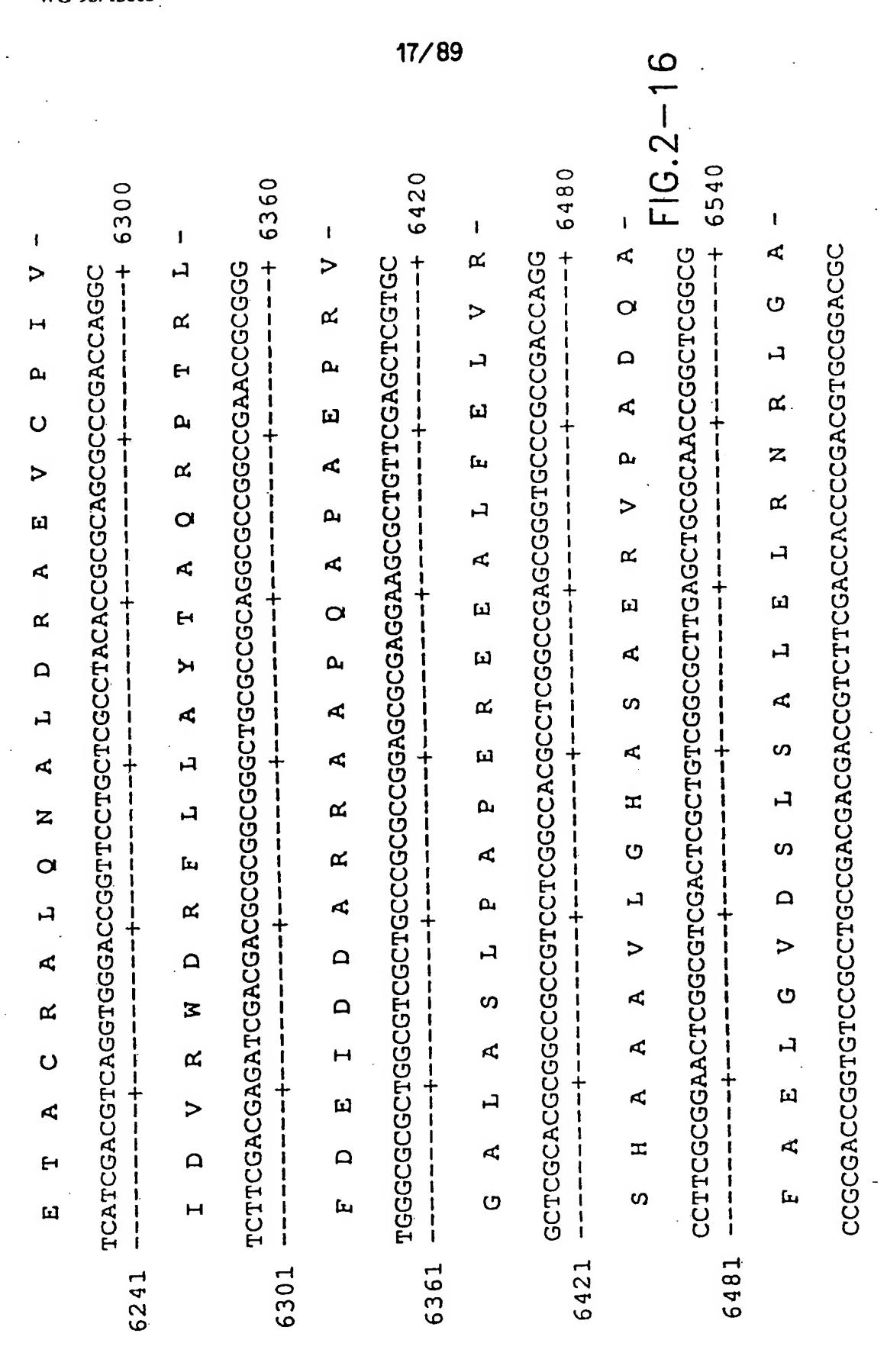
÷

					14/8	9			7	<u> </u>	
5220	·	5280	t <sup>-</sup>	5340	1	5400	į	5460	( (	FIG. 2- 5520	<b>.</b>
CAGGAGTGCTGTCCCTGCTGGTGGACGAAGCGGAGCGGA	LSLLAVDEAEPEEAPLAL-	TGGCTTCGCTGGCGACACGCTCAGCCTCGTGCAGGCGATGGTGTCGGCCGAACTCGGAT	LADTLS LVQAMVSAELGC-	GTCCGCTGTGGACGGAAAGCGCCGTCGCGACGGGGCCGTTCGAACGCGTCCGCA	WTVTESAVATGPFERVRN-	ACGCCCCCCACGCCCCTGTGGGCCTCGGGCGGTCATCGCGCTGGAGACCCCGCCG	HGALWGVGRVIALENPAV.	TGTGGGGGGGCGTCGACGTGCCCGGGGTCGGTCGCCGAGCTGGCCCGGCACCTCG	G L V D V P A G S V A E L A R H L A	CGGCGGTCGTGTCCGGCGCCCGGTGAGGACCAGCTCGCGCTGCGCGCCGACGGGTGT	VSGGAGEDQLARADGVY
CAGGAGTG	> 5	TGGCTTCG	S S	GTCCGCTC	ם	ACGCCGC	Ä		IJ Œ		> **
5161	·	5221		5281		5341		5401	·	5461	 

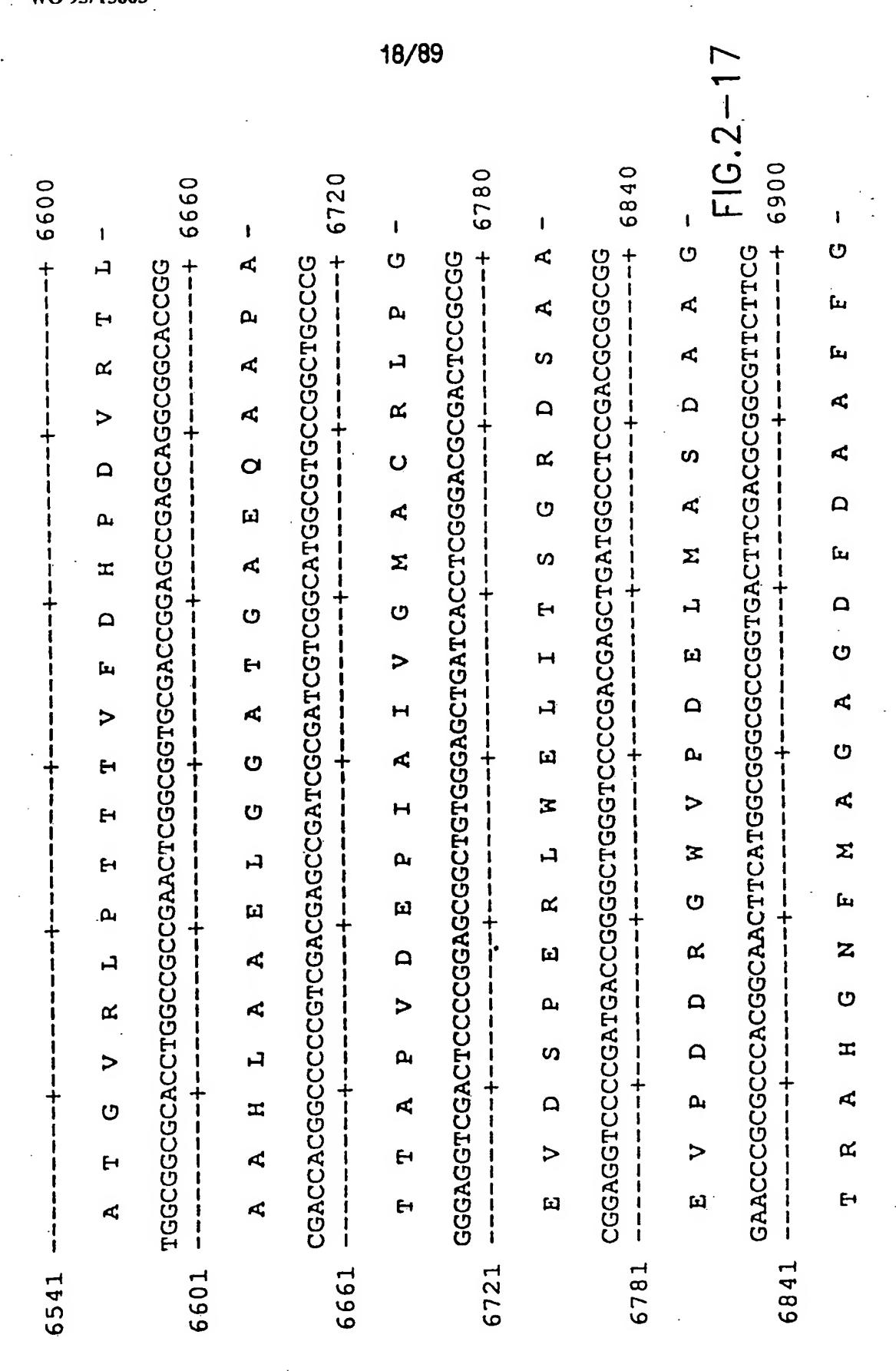


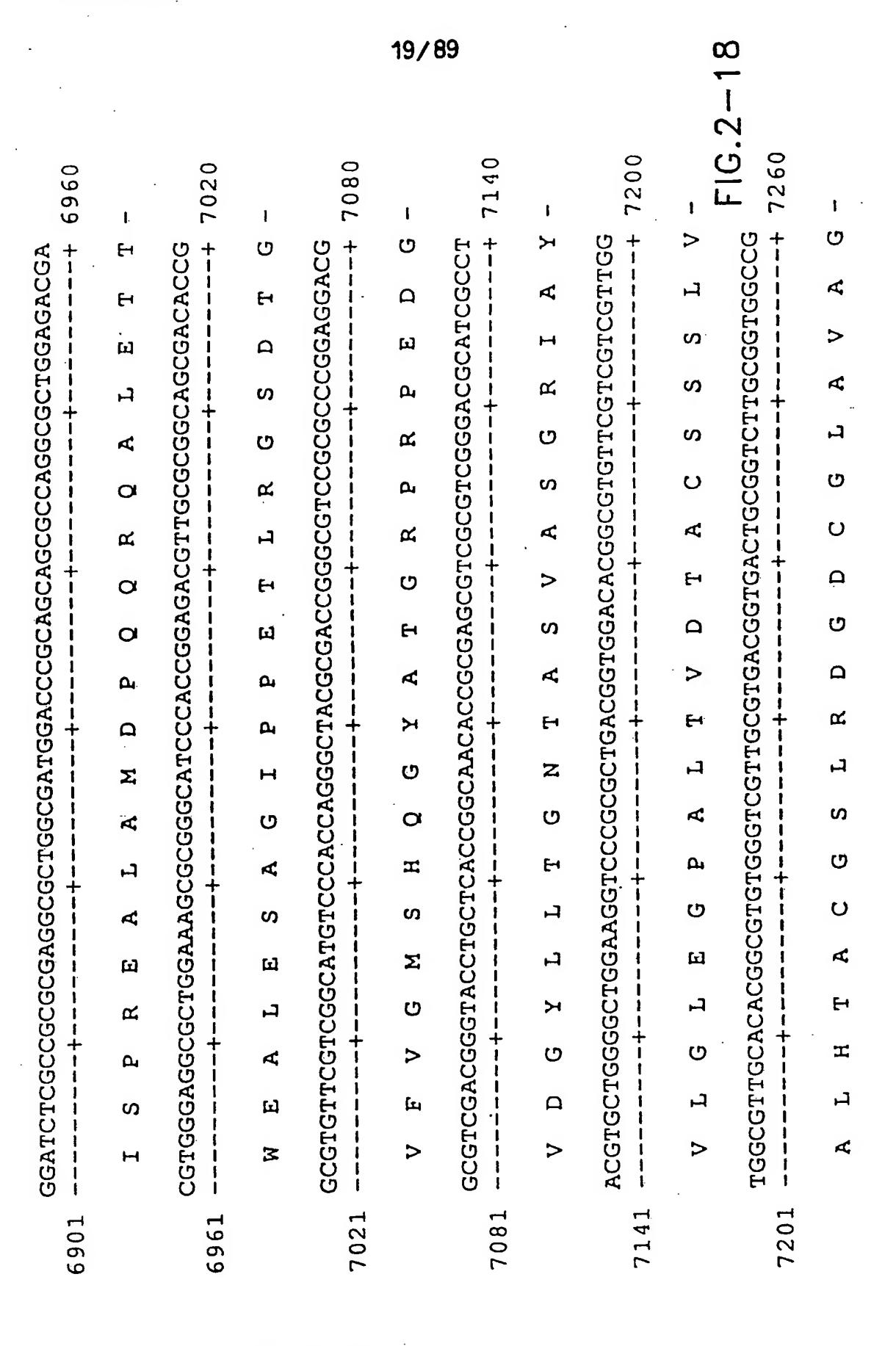


۲.



مطر





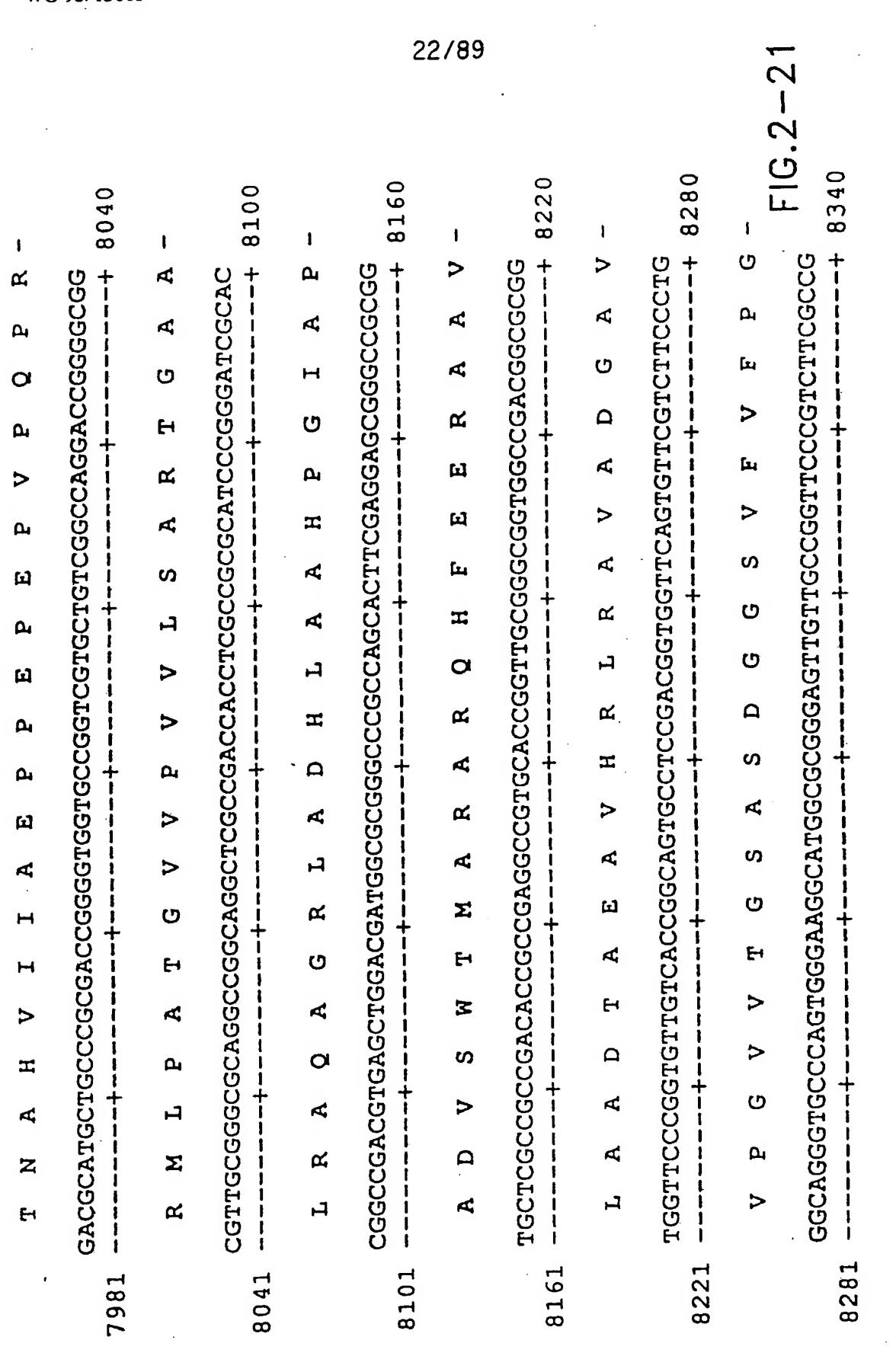
•					20/	89			, ,	6117	•
7320	ſ	7380	1	7440	1	7500	l	7560	ا ر	<b>F15</b>	I
GIGGIGIGICGGIGAGGICCGGAGGIGITCACCGAGITCICCCCCCCAGGGCGCGCGC	GVSVMAGPEVFTEFSRQGAL-	TCTCGCCGGACGCCCGGTGCAAGCCCTTCTCGGACGAGGCCGACGGATTCGGTCTCGGGG	SPDGRCKPFSDEADGFGLGE-	AGGGTTCGGCGTTCGTCCTCCAGCGGTTGTCCGACGCCAGGCGGGAGGGCCGCCGCG	GSAFVULQRLSDARREGRRV	TGCTCGGCGTGGCCGGGTCCGCGGTGAACCAGGACGGCGCGAGCAACGGGCTCTCCG	LGVVAGSAVNQDGASNGLSA	CICCGAGCGCGTCGCGCGGGTCAICCGCCGGGCGTGGGCGTGCGTGCGGGGATCA	PSGVAQQRVIRAWARAGIT	CGGGCGCGGATGTGGCGTGGAGGCGCATGGGACCGGTACGCGGCTGGGCGATCCGG	GADVAVERHGTGTRLGDPV
7261 -		7321		7381		7441	•	7501		75.61	

7621		. 08
	EASALLATYGKSRGSSGPVL-	
7681	TGCTGGGTTCGGTCGAACATCGGTCACGCGCGCGGCGGCGGGGGGGTGTCGCGGGCG	40
	LGSVKSNIGHAQAAAGVAGV-	
7741	TGATCAAGGTGCTGCTCGGCCTGGAACGCGGTGTGGTGCCCCCCCGATGCTGTGCCGGGGCG	0 0
	IKVLLGLERGVVPPMLCRGE-	21
7801	AGAGGTCGGGCCTCATCGACTGGTCCTCGGGCGAGATCGAGCTCGCAGACGGCGTGCGGG	/ <b>89</b>
	RSGLIDWSSGEIELADGVRE-	
7861	AGTGGTCGCCCCCCGCGGACGGGTGCGCCGGGCAGGTGTGTCGGCGTTCGGGGTGAGCG	920
	WSPADGVRRAGVSAFGVSG-	C
7921	GGACGAACGCGCACGTGATCATCGCCGAGCCCGCGGAACCGGGAGCCCGTGCCGCAACCGC++++++	. 1 <b>6.</b> 2 – 20

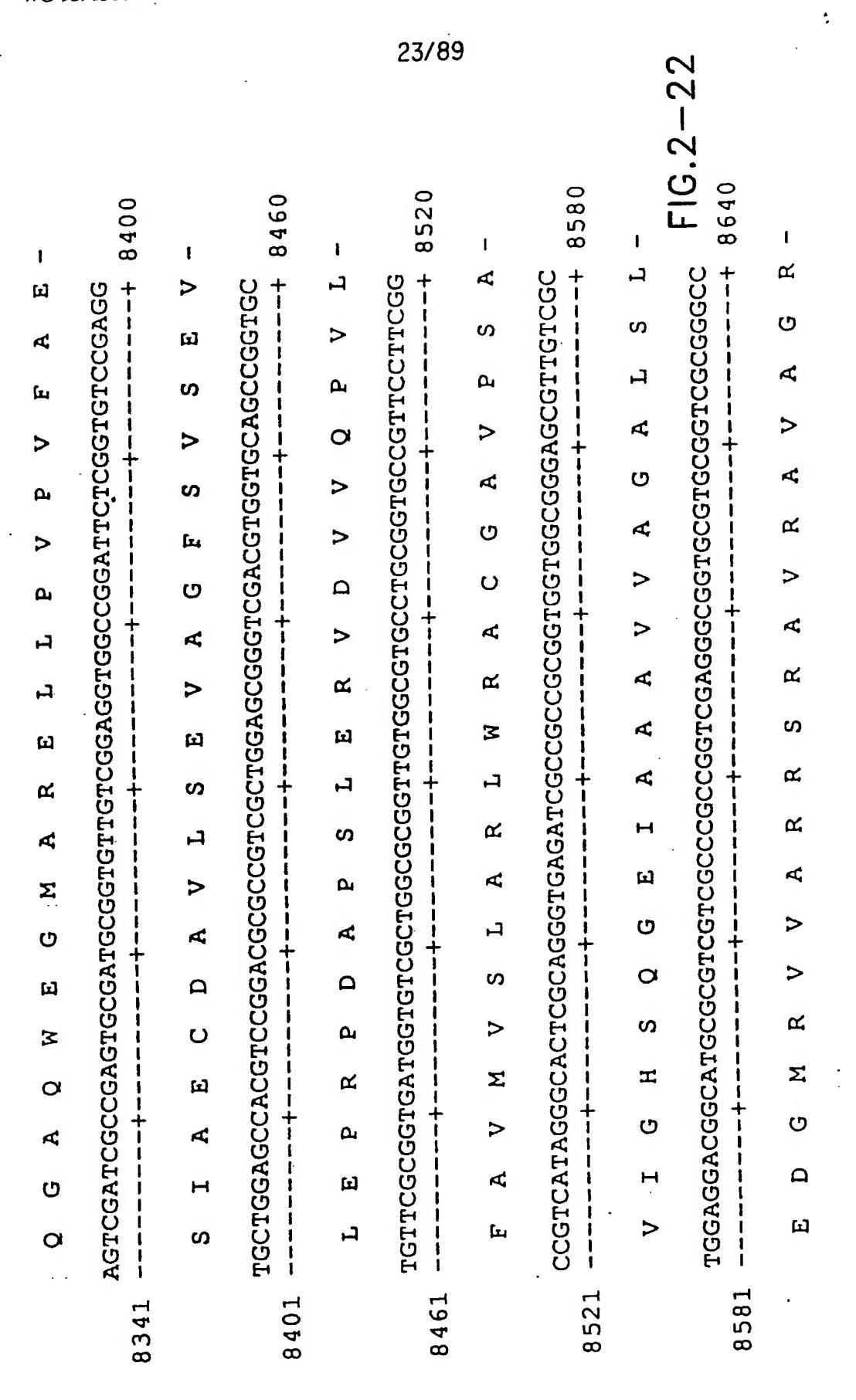
S.

क्र

æ



Á



Ē

•					24/8	39			<u>۱</u>	72	
007		. 09 /		820		8880		8940 <sub>©</sub>		F16.2-	•
GCGCGCGCCGCTCCGACGTCGAGAAGCTGCTCGCCGACG	RGGRSDVEKLLADD-	rggaggrcgcggrcaacggccccgacgccgrggrgg -+++++++ 87	EVAAV GPDAVVA-	SGCGCGAGTTCCTGGAGTACTGCGAGGGCGTGGGCATCCGCG	REFLEYCEGVGIRA-	ACTACGCCTCGCACGCGCACGTCGAGCCCGTGCGCGACG -++++++++	YASHTAHVEPVRDE-	CCGGGATCACCCCGCGACGGCCGAGGTGCCGTTCTTCTCCA	GITPRAEVPFFST-	TCGACGGCACCGAGCTGGACGCGGGCTACTGGTACCGCAACC	DGTELDAGYWYRNL-
GGGGGAGCATGCTCTCGGT	G S M L S V	ACAGCTGGACCGGCAGGCT	SWTGRL	CCGGTGACGCCCAGGCGGC	G D. A Q A A	CCCGCGCGATCCCGGTGGA	RAIPVD	AACTGGTCCAGGCGCTGGC	LVQALA	CCCTGACCGGCGACTTCC	LTGDFL
8641		8701		8761		8821		8881	•	8941	

					25/8	9			ر د	<b>47 -</b> 7
0906	1	9120	ı	9180	1	9240	ı	9300	_	9360
TGCGTCACCCGGTGGAGTTCCACTCCGCCGTGCAGGCGCTGACCGACC	R H	CGTTCATCGAGGTCAGCCCGCACCCGGTGCTGGCGTCGAGCGTCCAGGAGACCCTCGACG	FIEVSPHPVLASSVQETLDD	ACGCCGAGTCGCGCGGTGCTCGGGACGCTGGAACGCGACGCGGGCGACGCCGACC	AESDAVLGTLERDAGDAR	GCTTCCTCACGGCACTCGCCGACGCGCGCGCGCGGTCGCGGTCGACTGGGAAGCGG	FLTADAHTRGVAVDWEAV	TGCTCGGCCGGGCCGGACTGGTCGACCTGCCGGGTTATCCTTTCCAGGGCAAGCGGTTCT	LGRAGLVDLPGYPFQGKRFW	GGCTGCTGCCGGACCGCACCCCCTCGTGACGAGCTCGACGGCTGGTTCTACCGGGTCG
000		9061		9121		9181		9241		9301

\* 🐴

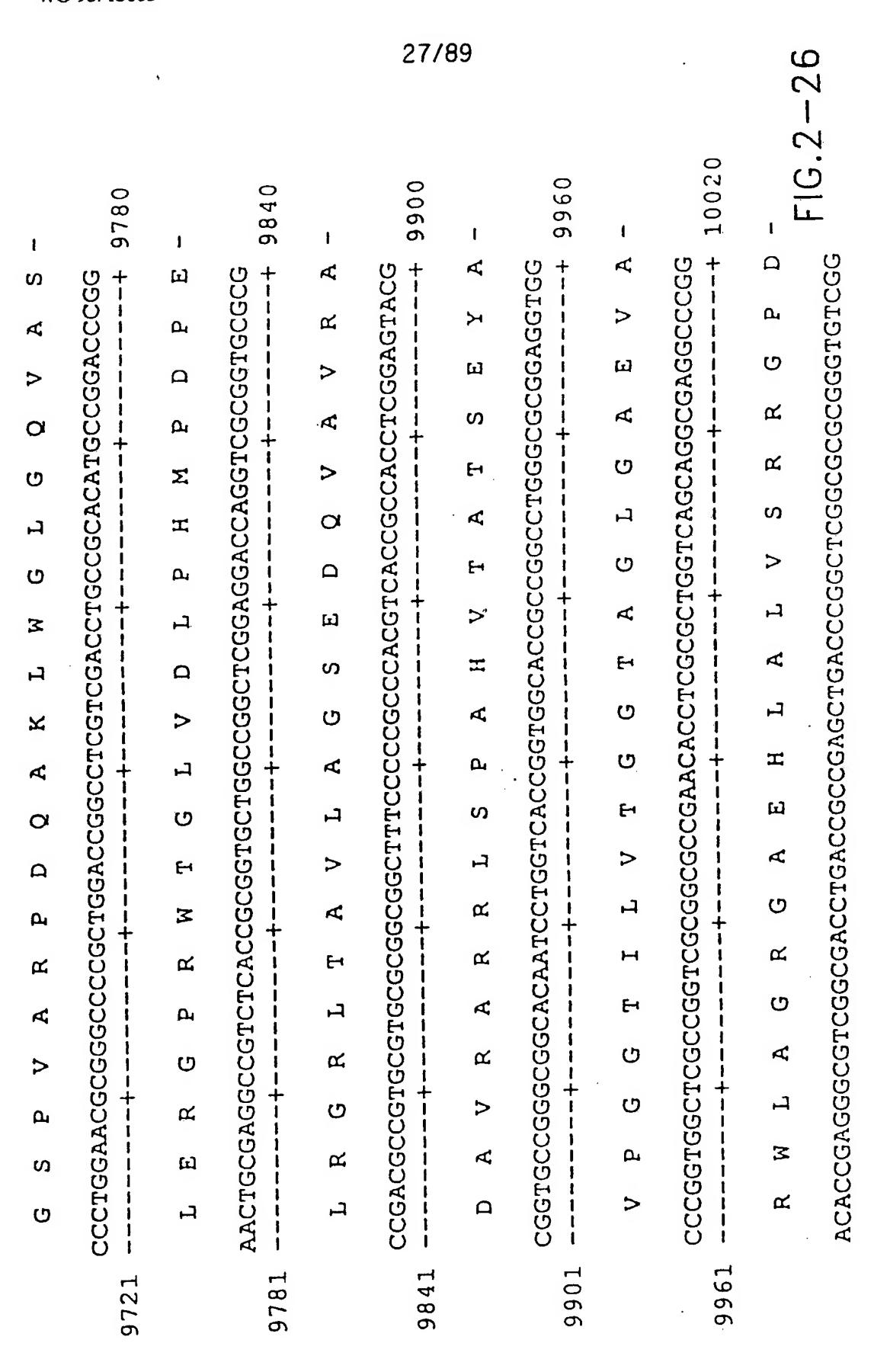
ŝ

\*

Þ

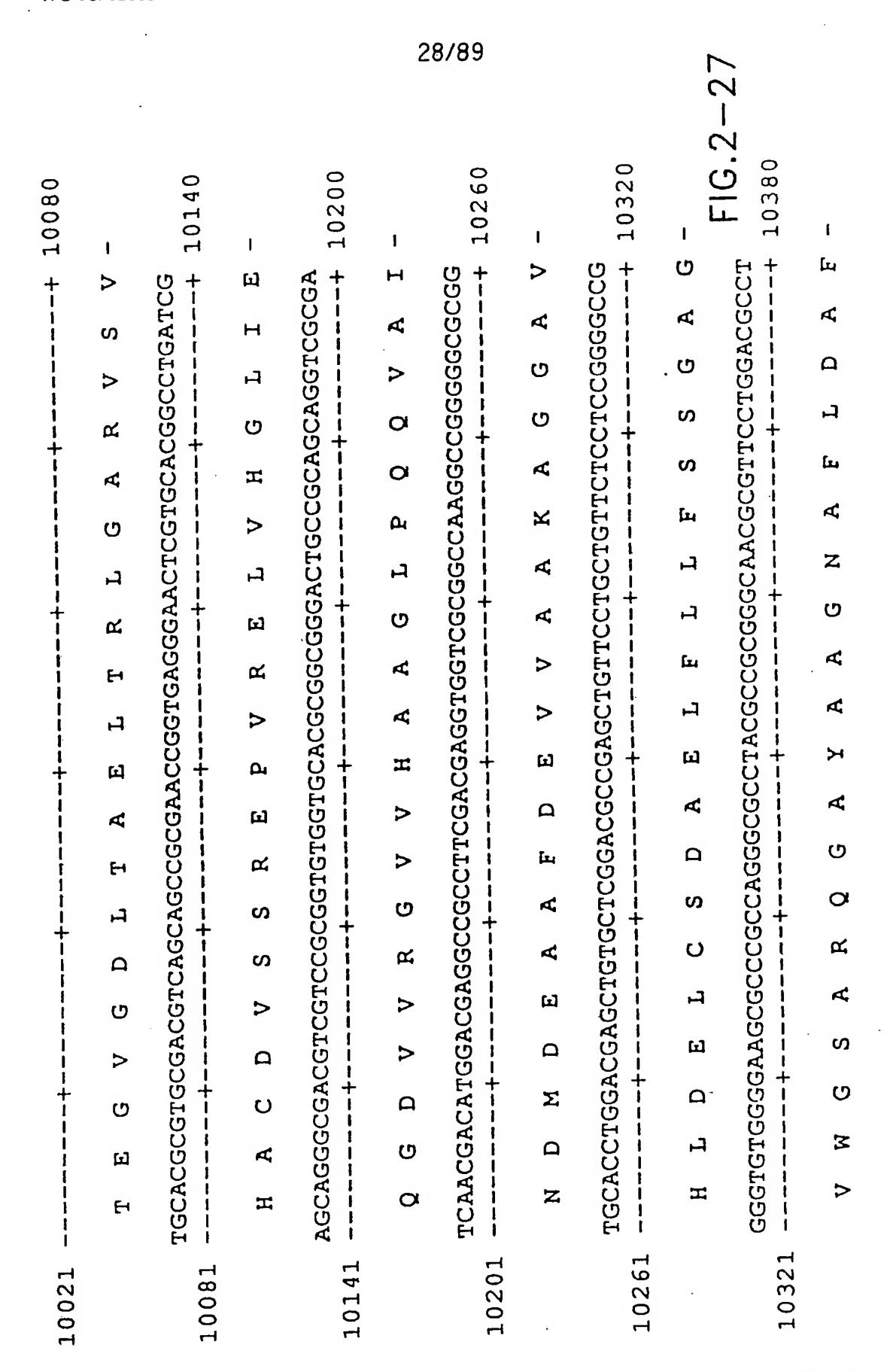
26/89 S 9720 9 0 9480 0 9420 0 96 9 95 g ŧ ١ GI K AACTCGACGCCGACGCCACGCCACTGTGGACGGTCACCTTCGGCGCGGTCGACG + 印 TGGTGTCGTTCGCCCTCGAGGGCGATGGTGCGGTGCAAACCCTTGTGCTGGTGCGGG CCGGCGCCGAACCGGGTCACGCGCGGCGTCGGCGGGGCTGGTCGGTGACTGCGGGGCG A CGCCGAGG TGCCGCGCTGCGCCCGTTGGCTCGTGGTGG Ŭ GIGGGGCIGGGCCAGGICG 区 G 回 K K K O H H GACCGTCGAGGTGCGGTCCGCGC H C K 3 T H بتأ U K S 3 H H K 9 G Q K Q CGGGCAGTCCGGTGGCCCGGGACCAGGCGAAGCT G 回 口 3 K C K 回 口 U . H K ACTGGACCGAGGTGCCGCGCTCCGAACC TGCCCGAGGGCACGAGGAGGACGGCTG Д Ω G 3 K Д 4 O  $\alpha$ G 团 μ Ω 回 E Ω S H Н H 回 CL. H U K 回 回 D<sub>4</sub> 只 团 **--+**-Д 二 > Д K 口 回  $\mathcal{O}$ 回  $\mathbf{p}_{\mathbf{q}}$ Ω S K 回 H G рц Н X 9601 1996 9541 9481 9421 9361

赤



÷

÷



ŕ

				•	2	9/89			· (	2-28	
10440	ı	10500	1	10560	ı	10620	ì	10680	1	F1G.	l
TCGCCCGGCACCGCGGGCCGCCTGCCCGCCACGTCGGTGGCGTGGGGGCTGTGGGG	A'R H R R G L P A T S V A W G L W A	CGGCGGCGCATGACCGGCGACGAGGAGGCCGTGTCGTTCCTGCGCGAGCGCGGTGTGC	AGGMTGDEEAVSFLRERGVR	GGGCGATGCCCGTACCGCCCTCGCCGCCCTGGACAGGGTGCTGGCTTCCGGGGAGA	AMPVPRALDRVLASGET	CGGCGGTGGTCGTGACGTGGACTGGCCCGCCTTCGCCGAGTCCTACACCGCCGCCC	AVVVTDVDWPAFAESYTARR	GGCCCCGGCCGTTGCTCGACCGCATCGTCACGACCGCGCCGAGCGGGCCGGGCCGGAAC	PRPLLDRIVTTAPSERAGEP	CGGAGACGGAGCCTGCGCGACCGGCTGGCGGGTCTGCCGCGTGCCGAGCGGACGGCGG	ETESLRDRLAGLPRAERTAE
10381		10441		10201		10561		10621		10681	

			٠		30	0/89			C	ה א	
10800	I	10860	l	10920	ĵ	10980	i	11040	, [	FIG. Z-	l
AGCTGGTGCGCCTGGTCCGCACCAGCACCGCGTGCTGGGCCACGACGACCGAAGG	LVRLSTATVLGHDDPKA	CGGTGCGCGCGACCACGCCGTTCAAGGAGCTCGGGTTCGACTCGCTGGCGGCCGTCCGGC	VRATTPFKELGFDSLAAVRL	TGCGCAACCTGCTCAACGCGGCCACCGGGCTCCGCCTGCCGTCGTCGTCGTTCGACCCTGACCCTTCGACCCTTCGACCCTTCGACCCTTTCGACCCTTCGACCCTTCGACCCTTTCGACCCTTCGACCCTTTTCGACCCTTTTCGACCCTTTTCGACCCTTTTTTTT	RNLLNAATGLRLPSTLVFDH	ACCCGAACGCCTCCGCGGTCGCCGGTTTCCTCGACGCCGAGCTCGGCACCGAGGTCCGGG	PNASAVAGFLDAELGTEVRG	GGGAGGCGCCCTCGCCGGGCTGGACGCGCTGGAAGGCGCCCTGCCGAGGTGC	EAPSALDALEGALPEVP	CCGCAACCGAGCGGAAGAGCTGGTACAGCGCTTGGAACGGATGCTCGCCGCGCTACGCC	A T E R E E L V Q R L E R M L A A L R P
10741	•	10801		10861		10921		10981	-	11041	

		31/89	20	) )	
11160	1219	12702		7 16.4- 12822	<b>J</b>
CGGTCGCCCAGGCCGCCGCCTCCGGGACCGGCCCAACCCGTCCGGCGACGACCTGG+++++++	GCGAGGCGGGCGTGGACGTCGAAGCACTCGGCCGGGAGCTCGACGGCGATTGA+++++++	CCGCCGATTGGAGAAAGGTGACTGACAGCGAGAAGGTGGCGGAGTACCTCCGTCGGGCG  3+++++++	ACGCTCGACCTGCGTGCCGCCGGCAGCTGCGAGCTGGAATCCGACCCGATCGCC  3+++++++	ATCGTCAGCATGGCCTGCCGCCTGCCGGGCGGGTGAACACCCCCGCAGCGGCTGTGG	I V S M A C R L P G G V N T P Q R L W E CTGCTGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
11101	11161	12643	12703	12763	

73		2		02				22	(	6.2-51 82	
1288	į	1294	į.	1300	t	130	ı	131		F 1(	1
+	REGGETLSGFPTDRGWDL-	GCGCGGCTGCACCACCCGGACAACCCCGGTACCAGCTACGTCGACAAGGGCGGG	L S Y U D G G G H H L	TICCICGACGACGCGGGGCTICGACGCGGAGTICTICGGCGTCTCGCCGCGCGCGAGGCC 13	DDAAGFDAEFFGVSPREA-	GCGGCCATGGACCCGCAGCGGCTGCTGCTGGAGCGAGCTGGGAGCTGGAGAAC 13	M D P Q Q R L L E T S W E L V E N -	GCCGGCATCGACCCGCACTCGCGCGCGTACCGCGACCGGCGTCTTCCTCGGAGTGGCG	IDPHSLRGTATGVFLGVA	AAGTTCGGCTACGGCGAGCACCGCCGCGGGGGACGTCGAGGGCTACTCGGTCACC 13	GYGEDTAAREDVEGYSVT
; ; ;	H	 BCBC	A R	TTCC	ᄪ	9909	A A	9229	S G	AAGT	자 [편
12823	, <b>4</b>	12883	·	12943		13003		13063		13123	

32/89

					33	3/89			7.0		
13242	l	13302	}	13362	l	13422	I	13482	ر راح	13542	Į
GGTGTGGCGCCCCCCCTCCGCCTCGCATCTCCTACCATCATGGCCTACCATGGCCTTCGTGCTCTTCTTACACCATGTTCTTACACCATGTTCTTACACCATGTTCTTACACCATGTTCTTACACCATGTTCTTACACCATGTTCTTACACCATGTTCTTACACCATGTTCTTTACACCATGTTCTTACACCATGTTCTTTACACACAC	GVAPPSGRISYTMGLEGP	TCGATCAGCGTCGACACCGCGTGCTCGTCGCTGCTGCTGCTGCTGCTGCGGTCGAG	SISVDTACSSLVALHLAVE	TCGCTGCGCAAGGGGGGTCGTCGATGGCGGTCGTCGTCGTGGTGCCGCGGTGATGGCGACC	S L R K G E S S M A V V G G A A V M A T	CCGGGGGGTGTTCGTCGACTTCAGCCGCGCGCGCGCCCCCCCC	PGVFVDFSRQRALAADGRSK	GCGTTCGGTGCCGGCCCGACGGGTTCGGCTTCTCCGAAGGCGTCACCCTGGTCCTGCTC	AFGADGFGFSEGVTLU	GAGCGGCTGTCGGAGGCGCGCGAAACGGGCACGAGGTGCTGGCGGGTGGTTCGCGGCTCG	ERLSEARRGHEVLAVVRGS
13183		13243		13303		13363		13423		13483	

Ġ

Ŧ

Ť

34/89

**C1** 2 390 **C1** 662 602  $\infty$ 3 ന Н  $\vdash$ -CGCAACGGGGAACTGCCGCGCGCCCTGCACGTCGAGGAGCCCCACGCCGCACGTCGACTGG ATCGGCCACACCCAGGGGGGGGGGGGCGTCACCGGCCTGAAGGTGGTCCTGGCGCTG GCCGCTCTGGCTCGGTGAAGTCCAAC CGACCCGATCGAGGCGAACGCGCTGCTGGAC AGGGTCATCCGGCCCCTCGAGAGCTGCGGTCTGGAGCCCGGCGACGTCGACGCGGTG 3 Z 民 Ω S K Q X 口 K I K Ωı **D**4 S Z Q 9 H 又 S A  $\mathcal{O}$ S  $\Omega_{\mathbf{1}}$ 回  $\Omega_{i}$ D<sub>4</sub> 团 Z H B 团 回 Ф S H **D**  $\Omega$ H ACCTACGGCGGCGACGCGGCCG GAGGCGCACGGCTACGGCGCTCGG 工 K C Ŋ H G Ω S Z K K 回 S æ K Ω H K D<sub>1</sub> A K G  $\mathfrak{O}$ H Ø Ω  $\vdash$ Q Ω 回 E 只 C PC, Q G 二 U 工 H Z Z Q R 又 H 团  $\alpha$ K 13843 13783 13723 13663 13603 13543

					35/	<b>'8</b> 9			7 7	1 0 4
13962	i	14022	1	14082	I	14142	1	14202	( ) L	r 16.2.
TCGTCCGGCGCGCGCGCTGCTGCCGGGCAACCAGCCGTGGCGGCGCGCGC	S S G G V A L L A G N Q P W R R G E R T	CGGCGCGCCCGTGTTTCCGGGATCAGCGGGACGAATGCGCACGTGATCGTCGAG	RRARVSAFGISGTNAHVIVE	GAAGCTCCTGAGCGCACCGGGAGACCACCGCGCACGACGGCCGACCGGTTCCGCTG	EAPERETTAHDGRPVPL	GIGGIGICCCCCCCCACGACGCCGTIGCGGCCCCAGAGCCCCCAGAICCCCGAGCTG	V V S A R T T A A L R A Q A A Q I A E L	CTCGAACGCCCGGACGCCTCGCCGGGGTCGGGCTGGGCCTGGCCACGACCCGCGCCC	LERPDADLAGIUGLATTRA	CGCCACGAGCACCGCGCCGCGTGGTGGCATCGACCCGCGAGGAAGCGGTGCGCGACTG
13903		13963		14023		14083		14143		14203

					30	6/89				, ,	0
ţ	14322	1	14382	ł	14442	1	14502	1	14562	. (	F16.
HEHRAAVVASTREEAVRGL	CGGGAGATCGCCGCCGCGCGACGCCGACGCCGTGGTCGAGGGCGTCACCGAGGTG	EIAAGAATADAVVEGVTEV	GACGGGCGCAACGTCGTTCCTGTTCCCGGGGCAGGGTTCGCAATGGGCCGGCATGGGT	GRNVVFLFPGQGSQWAAG	GCCGAGCTGCTGTCGTCGCCGGTGTTCGCCGGGAAGATCCGGGCCTGCGACGAGTCG	A E L L S S S P V F A G K I R A C D E S	ATGGCCCCGATGCAGGACTCTCCGACGTGCTGCGTCAGGCGCCGGGGGCGCCCG	MAPPULRQAPP	GGCCTGGACCGGGTGGTGCAGCCGGTGTTGTTCGCGGTGATGGTGTCGCTGGCG	GLDRVDVVQPVLFAVMVSLA	GAGCTGTGGCGCTCGTACGGCGTGGAGCCCGCGGCGGTCGTGGGGCACTCGCAGGGCGAG
<b>K</b>	CC 14263	R	14323 -	Ω	14383		14443	·	14503		14563

					37/8	39			<b>26</b>	
ı	14682	i	14742	<b>1</b>	14802	į	14862	1	FIG.2-	1
ELWRSYGVEPAAVVGHSQGE	ATCGCCGCCGCACGTCGCCGGGCGCTCACGTTGGAGGACGCGGCGAAGCTCGTCGTG	I A A H V A G A L T L E D A A K L V V	GGCCGCAGCCGCCTGATGCGGTCGCTCTCCGGGGAGGGCGGCATGGCCGCCGTCGCGCTG	GRSRLMS LSGEGGMAAVAL	GGCGAGGCCGCGGTGCGCGCCTGCGGCCGTGGCAGGACCGGCTCTCGGTGGCCGCG	GEAAVRERLPWQDRLSVAA	GTCAACGGTCCCCGGTCGTGGTCTCCGGCGAGCCCGGCGCGCTGCGGGCGTTTTTCC ++++++-	VNGPRSVVVSGEPGALRAFS	GAGGACTGCGCGGCGAGGGCATCCGCGTCGACATCGACGTGGACTACGCTCGCAC	EDCAAEGIRVRDIDVDYASH
	4623		14683		14743		14803		14863	

李

5

Ė

38/89

15282 15222 62 15042 02 982 S ហ 4  $\leftarrow$ CTCGGCCTTCCTGCGGTCGATGGCCACCGCG GGCTGACGGTGCCGAAGACGCGGGTCGTAGTC GTTCATCGAGGTCAGCCCGCATCCGGTTCGTG CIGGAIGCCCGGIACIGGIACCGCAACCIGCGAGACGGIGCGCTICGCCGACGCCGIG TATGGACGGCACCGAG CTCCTCGAAACGACCGGCGACATCGCGCCG A, 回 D4 H > H A. A Δι Q 9 H  $\Sigma$ K 工 K Q Ω S Ω ρι  $\Sigma$ [I U CGCCCGGCGCGGTGACGTTCCACTCCACTGTGGAGTCGCGGTC 只 团 S K S H 口 A > K H 9 H 回 H S 回 K H 团 回 Н S A [L 民 > 니 TCGCCGCAGATCGAGCGGGTCCGCGAGGAA GTCCAGGCCGTCGAGGAGGCGGTCGAAGA GGCTCGCTGCACCGCGGCGGTGACCT ACGCGGCTGGCGGAGTCGGGGATACGACGC H 回 H 딥 回 Z S 回 <mark>ර</mark> × 民 I · 以 111 G K C × [t4 > Ω 回 S H N, 1 回 团 × > 回 I A ፈ K H Н H Ø K O Ø S K ρι  $\Omega_{\mathbf{i}}$ U > H Н K S 15223 15163 15103 14923 15043 14983 SUBSTITUTE SHEET

CAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG						39	/89			7	χ γ
CAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	15342	1	15402	1	15462	1	2	1	.558		F <b>IG.</b> 5642
SUBSTITUTE SHEE	CACGTGTCCGGTGTGGACATCAGGTGGGACGTCGCTCTGCCCGGCGCCCCGCCCCGCCCCCCGCCCCCGCCCCCC	VSGVDIRWDVALPGAPF	CTGCCGACGTATCCGTTCCAGCGCAAGCGCTACTGGCTCCAGCCCGCCGCACCCGC	PTYPFQRKRYWLQPAPPA	GCCTCCGACGAGCTGGCCTACCGCGTTTCCTGGACTCCGATCGAAAGCC 5403+++++++	SDELAYRUSWTPIEKPES	AACCTGGACGGCGACTGGTTGTCACACCCCTCATCAGTCGGAGTGGACGGA	N L D G D W L V T P L I S P E W T E	CTGTGCGAGGCCATCAACGCCAACGGTGGCAGGGCGTTGCGCTGCGAGGT	LCEAINANGGRALRCEVDT	GCTTCGCGCACTGAGATGGCCCAGGCCGTCGCACAGGCCGGAACGGGATTCCGGGGCGT(

5

Ş

40/89 6002 2 2 5702  $\infty$  $\infty$ 57 S S S 1 --- $\leftarrow$ GGCGGCGGCCGAGGACCACCTCGCGGTCCGGGACGGCCGCCTCCACGGCCGTCGC GGCGCTGCTCTCCTCGTGGCAGTCCTGCGC GCTGGAGCTGCCGGGCCGCTGGGGGGGGTGTG CGACGACGTCGCCCGGCCTGCGCAGACC GGCCTGCTCACCCTGGTCCTGGGCGATGCCGGGGTCGACGCACCGGTGTGGTGC CTCTCGTTGCTGTCGGACGAATCCGCCTGCCGTCCGGGGGGTTCCTGCCGGTGCGGTC **M** H O> O O 3 K C C A G 只 K 3 PI K Д म्य 区 只 K Ω 9 H C K Q H H  $\Omega_{\mathbf{i}}$ Н > U 0 只  $\Omega$ 4 回 A K 只 K **山**. D  $\mathbf{O}$ Q K ACCGCGCACGGCTTCGCGCAGGTCGCCGG GTCGACCTGCCCGAATCGGTCGACGACGC CTGACCCAGGGTGCGTCCGCACTCCCGC G C K **A** <u>ρ</u> S Ω > H 团 Ø 1---+-> Q K Q K Ω S A > S Σ 山 Ŀц K Н S 回  $\Omega_{4}$ C  $\mathcal{O}$ H Н H 工 O H 工 H S S 1 >  $\vdash$ Н G Н K 15943 15883 15823 15763 15703 15643

G G G R A E D H L A V R D G R L H G R R - GTCGTCGCGCAGTCCGCAGTCCGCGAGCTGGAGCTCGGAGCTGGAGCCCGCAGGGAGCTGGAGCCCGCAGGGAGCTGGAGCCCGCAGGCCGTGGAGCCCGCAGCCGCGAGCCGGAGCCGGAGCCCGCAGCCGCAGCCGAGCCGAGCCGGAGCCGGAGCCCGCAGCCGCAGCCGAGCCGGAGCCCGCAGCCCGCAGCCGAGCCCGCAGCCGAGCCGAGCCCGCAGCCCGCAGCCGAGCCCGCAGCCCCGCAGCCCCGCAGCCCCGCAGCCCCGCAGCCCCGCAGCCCGCAGCCCGCAGCCCGCAGCCCCGCAGCCCCGCAGCCCCCC	
	6062
RASLPQSGSRSWTPHGTV-	
CTGGTCACCGGCGCGAGCCCCGTCGGCGACCAACTGGTGCGGGGTTCGCTCGC	6122
TGAASPVGDQLVRWLADR-	
GGAGCCGAGCGGCTGGTGCTGGCCGGAGCCTGTCCGGGCGACGACCTGCTGGCCGCGGTC	6182
ERLVLAGPGDDLLAAV -	/89
GAGGAAGCGGCGCATCGTGTGCGCCCAGGACGCGCGCGCG	.6242
AGASAVVCAQDAALREA -	
CTCGGCGACGAGCCGGTGACCGCTCGTGCACGCCGGAACCTTCGGCAGC++++	F16.2-40
DEPVTALVHAG.TLTNFGS-	•
ATCAGCGAAGTCGCAGGAGTTCGCCGAGACGATCGCGGCCAAGACCGCGTTGCTC	

					4	12/89			7	  -  -	
16362	ľ	16422	1	16482	í	16542	1	16602		16662	
	I S E V A P E E F A E T I A A K T A L L	GCCGTGCTGGACGAAGTCCTCGGCGACCGGCCGTCGAGCGGGAGGTCTACTGCTCGTCG	AVLDEVLGDRAVERFVYCSS	GTCGCCGGGATCTGGGGCGGCGCCGGATGGCCGCCTACGCGGCAGGCA	V A G I W G G A A A A A A Y L	GACGCGCTGGCCGACCACCGCGCGCGGGCCCGCTCGTGCACCTCGGTCGCCTGGACG	DALAEHHRARGRSCTSVAWT	CCGTGGGCGCTGCCGGGGGGGGGGTGGACGACGGCTACCTGCGGGAACGCGGACTGCGC	PWALPGGAVDDGYLRERGLR	AGCCTCTCCGCCGACAGGGCGATGCGCACCTGGGAGCGGGTGCTGGCCGCCGCCGGCCG	SLSADRAMETWERVLAAGPV
16303		16363		16423		16483		16543		16603	

					43	3 /89			~	1 4 7	
16722	i	16782	1	16842	1	16902	į	16962	ן ה	17022	
TCGGTCGCGGTGGCCGACGTGCCGGTGCTCAGCGAAGGCTTCGCCGAAGGCTTTCGCCGACGCACGCGACGCGACGCGACGCGACGCGACGCGACGCGACGCGACGCGACGCGACGCGACGCGACGCACGCGACGCGACGCGACGCGACGCGACGCGACGCGACGCGACGCGACGCGACGCGACGCGACGCACACGCGACGCGACGCGACGCACACGCGACGCGACGCACACGCGACGCACACGCGACGCA	S V A V A D W P V L S E G F A A T R	CCGACCGCGCTGTTCGCCGAACTCGCCGGCCGCGGGGGGGG	EH .	AGCGGACCGACCGGCCGGCACAACGGCTCGCGGGGCTTTCCCCGGACGAGCAGCAG	SGPTGEPAQRLAGLSPDEQQ	GAAAACCTGCTCGAACTCGCGAACGCGGTTGCCGAGGTGCTTGGCCACGAGTCCGCC	ENLLELVANA VAEVLGHESA	GCCGAGATCAACGTGCGCGCGCGTTCAGCGAGCTCGGACTCGACTCGCTCAACGCGATG	A E I N V R R A F S E L G L D S L N A M	GCCCTGCGCAAGCGCCTGTCGGCGAGCACCGGCCTGCGGGCTGCCGGCGTCGCTGTT(	ALRKRLSASTGLRLPASLVF
16663		16723	) i .	16783		16843		16903		16963	

•		44	1/89	777	<b>)</b>
17082	17142	17202	17262	17322 - FIG 2	173
GACCACCCCACCGTCGCGCAGCACCTGCGCGCCCGGCTCGTCGGTGACGCC  ACCACCCCCCCCCC	GACCAGGCCGCGCGCGCGCGCGGCCGACGAGCCCATCGCCATCGTC  3+++++++	GGCATCGGCTGCCGTTTCCCCGGCGCATCGGCTCGCCCGAGCAGTTGTGGCGGGTGCTG  3++++++++	GCCGAGGGCGCGAACCTCACCACCGGCTTCCCGGCCCGACCGGGGCTGGGACATCGGGCGGG	CTCTACCACCCGGACCACCCCGGCACCAGCTACGTGGACAAGGGCGGGTTCCTC	ACCGACGCGGATTTCGACCCGGGCTTCTTCGGCATCACGCCCCGCGAAGCGCTGGCG  3+++++++++-
17023	17083	17143	17203	17263	1732

ż.

3

					45/	89				† †
17442		17502	i	17562	1	17622	ł	17682		17742
ATGGACCCGCAGCGCCTCATGCTGGAGACGGCGTGGAGGCAGTGGAACGCGGGC	M D P Q Q R L M L E T A W E A V E R A G	ATCGACCCCGACGCCCCGACCGACACCGGCGTCTTCGTCGGCATGAACGGCCAG	I D P D A L R G T D T G V F V G M N G Q	TCCTACATGCAGCTGCTGGTGAGGCCGAACGCGTCGACGGCTACCAGGGCCTCGGA	SYMQLL'AGERERVDGYQ	AACTCCGCGAGCGTGCTCTCCGGGCGCATCGCCTACACCTTCGGCTGGGAGGGCCCGGCG	NSAS; VLSGRIAYTFGWEGPA	CTGACGGTGGACACCGCGTGCTCGTCGTCGGTCGGCATCCACCTCGCGATGCAGGCG	LTVDTACSSLVGIHLAMQA	CIGCGCCCCGTGAGIGCICCCIGCCGCTGGCCGGCGCGTCACCGGTCATGICCGACCCG
17383		17443		17503		17563		17623		17683

46/89 8042 17982 922 62 17802 œ į GTGATCCGGCAGGCGCCCCCCCTCGGGCGTGCCGGCCGGGACGTCGACGTCGTGGAG GTCAACCAGGACGGTGCAACGGTCTCGCCGCTCCCAACGGCCCGTCGCAGGAGCGG GTCGGAAGGCGTCGCCGCGCTGGTGCTGGAG CGGGCTCGCTCCGACGGTCGCTGCAAGGCG 回 又 K 回 Д 回 S H X A > Q C O S Q S K H K  $\Sigma$ > Д H K U Ω Q. K D H K Z K S > K  $\Omega_1$ U K C р A 回 C > 4 Q S G K G TACACCTTCGTCGACTTCAGCACGCAGCG TTCTCCGCGGGCCGACGGCTTCGCGCT 工 K S  $\mathcal{O}$ G K OK A, Z Z [tı H H K S K **O** S S Н A, PL, Ŀı  $\mathbf{O}$ K  $\mathcal{O}$ A  $\Box$ 回 O Ω Ľ >  $\mathcal{O}$ 只 Q S K 只 H Z S H  $\alpha$ > Ω<sub>4</sub> 7 H 17983 17923 17863 17803 17743

SUBSTITUTE SHEET

ŝ

う

	•				4	7/89			7	† 	
18102	1	18162	Į.	18222	t	18282	1	18342		18402	ı
GCGCACGGGACGGCACCGACCTCGGCCGATCGAGGCCGGCGCGCTCATCGCACC	AHGTELGDPIEAGALIAT	TACGGCCAGGACCGCCGCCGCTGCGGCTCGGCTGAAGACCAACATCGGCCAC	YGQDRPLRLGSVKTNIGH	ACCCAGGCCGCGGCGCGCGCGTGATCAAGGTCGTGCTGGCGATGCGGCGGCGGGGGGGG	T Q A A G V I K V L A M R H G	ATGCTGCCCGGTCGTTGCACGCCGACGAGCTGTCCCCGCACATCGACTGGGAGTCGGGG	M L P R S L H A D E L S P H I D W E S G	GCCGTGGAGGTGCTGCGCGAGGTGCCGTGGCCGGCGGGTGAGCGCCCCCCGGCGGCG	AVEVLREEVPWPAGERPRA	GGGGTGTCGTCCTTCGGCGTCAGCGGAACCAACGCGCACGTGATCGTCGAAGAGGCACCA	GVSSFGVSGTNAHVIVEERAP
18043	Æ ş	18103		18163		18223		18283		18343	

ŝ

5

					48	/89		,	7	/ +	
18462	l	18522	1	18582	.1	18642	l	18702	_	F16.	
GCAGAGCAGGCCGCCCCGCACCGAGCGCGGTCCGCTGCCGTTCGTGCTGTCTCGTGTCCGC	EQEAARTERGPLPFVLSGR	AGCGAAGCCGTGGCCCAGGCCCGCGCGCTCGCCGAGCACCTGCGCGACACCCGG	EAVVARRALAEHLRDTP	GAGCTCGGCCTGACCGCGCGTGGACGCTCGCGACCGGCAGGGCGCGCGGTTCGACGTG	LGLTDAWTLATGRARFDV	CGAGCCGCCGTGCTCGCCGACCGCGCGCGCGTGTGCGCGGGAGCTGGACGCGCGCCCCCCCC	RAAVLGDDRAGVCAELDALA	GAGGCCGCCCGTCGCCGTCGCGCCGGTGACCTCCGCGCGCG	ග	CTGGTCTTCCCCGGCCAGGGCGCGCAGTGGGTCGGCATGGCACGCGATCTGCTGGAATCC	о . Б
G 18403 -	K	A 18463 -	ഗ	G 18523 -	EI EI	18583 -		, ה ה ה	) ! 0	, C	7 0 7 0 7

ぅ

Ģ

					49/8	19			0 7	7 - 40
18822	ı	18882	ı	18942	ı	19002	1	19062	C L	19122
TCCGAGGTGTTCGCCGAGTCGATGAGCCGGTGCGCCGCGCGCTCTCGCCGCACGGACGACGACGACGACGACGACGACGACGACG	SEVFAESMCAEALSPHTD	TGGAAGTTGCTCGACGTCCGCGCGCGCGGCGGTCCCGACCCGCACGAGCGCGTCGAC	WKLLDVVRGDGGPDPHERVD	GIGCICCAGCCGGIGCICTICICGAICAIGGICICGCTGGCCGAGCIGIGGCGCGCACACACACACACACACACACAC	VLQPVLFSIMVSLAELWRAH	GGCGTGACCCCGGCCGC	G V T P A A V V G H S Q G E I A A H V	GCGGGCGCGCTGTCGCTGGAAGCCGCCGCGAAGGTGGTGGCCCTGCGCAGCCAGGTGTTG	AGALSLEAARKVVALRSQVL	CGCGAGCTCGACGACCAGGGCGGCATGGTCGGTCGGCGCGCGTCCCGCGACGAGCTGGAG
18763		18823		18883		18943		19003		19063

ACCGTGCTGGGGCGGGCGGGTGGCGGCGGGGGCTGGAACGGGGCTGGAACGGGGCTGGCACC  19123						50/	89				70	† 
19123	1	91	ì	92		6	I	93	i	942	ر الا	1948
	ELDDQGGMVSVGASRDEL	ACCGTGCTCGCGCGCTGGGACGGCCGTGTCGCGGTGGCCGCCGTGAACGGGCCTGGCAC	VLARWDGRVAAVNGPG	AGCGTCGTTGCCGGCCGACCGCGGAGCTGGACGAGTTCTTCGCCGAGGCCGAGGCGCGG	V V A G P T A E L D' E F F A E A	GAGATGAAGCCGCGCGGATCGCCGTGCGCTACGCCTCCCACTCCCCC	M K P R R I A V R Y A S H S P E V A	ATCGAGGACCGGCTCGCGGCCGAGCTGGGCACCATCACCGCCGTGCGGGGCTCGGTGCC	EDRLAAELGTITAVRGSV	CTGCACTCCACGGTGACCGGCGAGGTCATCGACACCTCCGCGATGGACGCCTCCTACTG	H S T V T G E V I D T S A M D A S · Y	CAGTGCTCTTCGAGCAGGCGGTGCGCGGGTCTGGT
		<b>1-1</b>			SUBS		SHE					

Ę

رش

						51/8	39		-50		
1	19542	1	19602	I	19662	1	19722	i	FIG.2-	I	
YRNLRRPVLFEQAVRGLVEQ	GGCTTCGACACCTTCGTCGAGGTGACCCCGCACCCGGTGCTGCTGATGGCGGTCGAGGAG	G F D T F V E V S P H P V L L M. A V E E	ACCGCCGAGCACGCGGGGCGCGAAGTCACCTGCGTGCCGACGCTGCGCCGCGCGAGCAGAGCAGAGC	TAEHAGAEVTCVPTLRREQS	GGACCGCACGAGTTCCTGCGAACCTGCTGCGGGCTCACGTGCACGGCGTCGGCGCGCGC	GPHEFLRNLLRRAHVHGVGAD	CIGCGICCGGCGGCCGGGGACGGCCGGCCGAGCIGCCCACCIACCCGTICGAACAC	LRPAGGRPAELPTYPFEH	CAGCGCTTCTGGCCGCCGCCCGCCCCCCCGCCGACGTCTCGGCGCTGGGCGTGCGCGCGC	QRFWPRPADVSALGVRG	GCGGAGCACCCGCTGCTCGCCGCGGTCGACGTGCCGGGCCACGGCGGTGCGGTGTTT
•	19483	~	19543		19603		19663		19723	•	

E

						52 /	89		L	_ ი	
19842	•	19902	1	19962		20022	1	20082		F16.2-	1
+	A E H P L L L A A V D V P G H G G A V F	ACCGGAAGGCTTTCCACCGACGAGCAGCCGTGGCTGGCCGAACACGTCGTGGCGGCCGG	TGRLSTDEQPWLAEHVVGGR	ACGCTGGTGCCGGGCAGCGTCCTGGTCTCGCGCTCGCCGGGGTGAGGACGTCGGG	TLVPGSVLVDLARAGEDVG	CTGCCGGTCCTGGAGGAACTGGTGTTGCAACGGCCGCTGGTGCTGGCCGGGGGGGG	LPVLEELVLORPLVLAGA	CTGCTGCGCATGTCGGTCGGCGCCCCAACGAGTCGGGGCGGCGGCGACGATCGACGTCCAC	LLRMSVGAPDESGRRTIDVH	GCCGCCGAAGACGTGGCCGACCTCGCCGACGCGCAGTGGTCGCAGCACGCCACCGGGACG	A A E D V A D L A D W S Q H A T G T
19783		19843		19903		19963		20023	•	20083	

						53/89			П С	70	
20202	1	20262	•	20322	1	20382	t	20442		•	
.CTCGCGCAGGGCGTCGCGGGTCCGAGGGATACCGAGCAGTGGCCGCCGGCGGAGGACGCC	LAQGVAAGPRDTEQWPPEDA	GTCCGCATCCCGCTCGACCACTACGACGGCCTCGCCGAGCAGGCTACGAGTACGGA	VRIPLDDHYDGLAEQGYEYG	CCGTCGTTCCAGGCCTGCGTGGCGCAAGGACGACTCGGTCTACGCCGAGGTG	PSFQALRAWRKDDSVYAEV	TCCATCGCGGCGGACGAGGTTACGCGTTCCACCCGGTGCTGCTCGACGCCGTGGCG	SIAADEEGYAFHPVLLDAVA	CAGACGCTCAGCCTGGCCCCTCGGCGAGCCGGGGGGGGAAAGCTGCCGTTCGCGTGG	OTLSLGALGEPGGKLPFAW	_	HILVEN
		20203		20263		20323		20383	) ) )		5 4 4 0 7

ŝ

3

					54/	89			ប	CC-7
20562	Į	20622	I	20682	1	20742	Ī	20802		F16.
GGGGCGGACGCGATGGCCCTGCGGGTCACCGACCCGGCAGGCCACCTGGTCGCCACGGTC +++++++	GADAMALRVTDPAGHLVATV	GACTCGCTGGTCGTCGCAGCACCGGGAGAAGTGGGAGCAGCCCGAACCGCGCGGTGGC	DSLVVRSTGEKWEQPEPRGG	GAGGGCGAGCTGCACGCTCTGGACTGGGGACGGCTAGCCGAGCCCGGCTCGACCGGTCGT	EGELHALDWGRLAEPGSTGR	GIGGICGCCGAIGCCICGGACCICGACGCCGICCIGGGICCGGIGAACCCCGAACCCC	VVAAADLDAVLRSGEP	GACGCGGTCCTGGTACGAACCCGAACGCGACGACCCCCCGCGCGCG	DAVLVRYEPEGDDPRAARH	GGCGTCCTCTGGGCCGCCGCTCGTGCGCCGCTGGCTCGAACAGGAGGAGCTGCCGGGC 03+++++++++
20503		20563		20623		2068		207		208

う

5

55/89 **C1** 62 **C1** 2 7 **C1** 22 0  $\infty$ **C1** 9 0 2 0 2 **C1** 2 ~ 2 99 GCCGCCTCTCGCCGCTCGCACCT TGCGGTTCCGGACAAC GCCGGAC  $\mathcal{O}$ Д Z IC C GT  $\alpha$ D O AG A, P S S Н TGTGCGCAGGCCGAG TCCGACGACGAC C  $\Box$ 回 Ω 回 S 4 A [1] GGGATGCTGCC 9 Q Ф Q Д X A S 回  $\Box$ O S Ū AC S C 民 Ü H 3 GCGACGCTGGTCATCGCCACGTCCGGCGGGTC GAACCCGGCGCCGCGATGTGGGGGCGTGAT K AC H P H K CCGCAGCTCGCGTTGCGCGGCGACGACGT CGGTTCGTCCTCGACACCGACGCGGA > 回 K Ľ H C U G Ω 1111+1 Ω 3 S Н ---+---B G H Σ  $\vdash$ B K Ω B K 1111111 П 口 B H K EH Н K 3 r K Q Гц  $\Omega_{\mathbf{i}}$ H S O1 K 回 K 9 21163 21103 21043 20983 20923 20863

						56/8	9			Ŋ	) () 
ı	21282		21342	1	21402	1	21462	1	21522	( ) ( )	716.4 21582
SVAFEPAPDVEQPLRAGEVR	GTGGACGTGCGCCACCGGAGTCAACTTCCGCGACGTCCTCCTCGCACTCGCATGTAT +++++++	V D V R A T G V N F R D V L L A L G M Y	CCGCAGAAGGCGGACATGGGCACCGAGGCCGCCGGTCGTCACGGCGGTCGGACCGGAC	P Q K A D M G T E A A G V V T A V G P D	GTGGACGCCTTCGCCCGGGAGACCGGGTGCTCGGCCTGTTCCAGGGAGCCTTCGCGCCG	VDAFAPGDRVLGLFQGAFAP	ATCGCGGTCACCGATCACCGGCTCCTCGCACGAGTGCCGGACGGCTGGAGCGACGCCGAC	I A V T D H R L L A R V P D G W S D A D	GCCGCGCCGTGCCCATCGCCTACACCACGGCGCATTACGCGCTGCACGATCTCGCGGGGG	A A A V P I A Y T T A H Y A L H D L A G	CTGCGCGCGGGTCAGTCGTGCTCATCCACGCAGCGGCAGGCGGTGTCGGCATGGCGGCC ++++++-
	21223		21283		21343		21403		21463		21523

					57/	89			ر م	
į	21642	1	21702	t	21762	ı	21822		F1G.2-	i
LRAGQSVLIHAAAGGVGMAA	GTCGCGCTGGCCCGCGGGGGGGGGGTGTTGGCCACCGCCGGCCG	VALARAGAEVLATAGPAKH	GGGACGCTGCGGCGCTCGGTCTCGACGAGCACATCGCTTCCTCCCGGGAGACCGGT	GTLRALGLDDEHIASSRETG	TTCGCCCGGAAGTTCCGGGAGCGCACCGGAGGCCGCGCGCG	FARRERTGGRGVDVVLNS	CTCACCGGGGAACTGCTCGAGTCCGCGGATCTGCTCGCCGAGGACGGCGTCTTCGTC	LTGELLDESADLLAEDGVFV	GAGATGGGCAAGACCGACCTGCGGGACGCCGGGGACTTCCGGGGCCGATACGCCCGTTC	EMGKTDLRDAGDFRGRYAPF
	21583		21643		21703		21763		21823	1 

 $\infty$ 

						59	/89				Ω Ω Ι
I	22302	i	22362	1	22422	1	22482	ŧ	22542		F1G. Z. 22.602
RGADAPGSDELRAEIEDLGA	TCCGCGGAGATCGCGGCTTGCGACACCGCCGACGCGCGCG	SAEIAACDTADALSALD	GGGCTGCCCCGGCCGCTGACCGGTGTCGTGCACGCGGGGGGGTGTGCTGGCCGACGGGCTG	G L P R P L T G V W H A A G V L A D G L	GTCACCTCCATCGACGAGCGGGGGTGGAGCTGCTGCGCGCCAAGGTCGACGCGGCG	V T S I D E P A V E Q V L R A K V D A A	TGGAACCTGCACGAGCTGACCGCGAACACCGGTCTGAGCTTCTTCGTGCTGTTCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCCTC	WNLHELTANTGLSFFVLFSS	GCGGCGTCGGTGCTAGCCGGCCCCGGGGCAGGGCGTGTACGCGGCCGCGCAACGAGTCGCTC	A A S V L A G P G Q G V Y A A N E S L	AACGCGCTGGCTGCCCTCCGGAGGACGCGGGCCTTCCCGCGAAGGCGCTCGGATGGGGA
	22243		22303		22363		22423		22483	•	22543

					a	60/8	9			_	ת ח 1
1	22662	1	22722	į	22782	I	22842	l	22902		716 22962
NALARTRGLPAKALGWG	CTGTGGGCGCAGCCAGCGAGATGACCAGCGGACTCGGCGACCGCATCGCCCGGACCGGGGGGGG	LWAQASEMTSGLGDRIARTG	GTCGCCGCCCTGCCGACCGAGCGGCGCTCGCACTGTTCGACAGCGCCCTGCGCCGCGGC	VAALERALFDSALRG	GGTGAGGTCGTGTTCCCGTTCAACCGTTCCGCGCTGCGCAGGGCCGAGTTCGTG	GEVVFPLSINRSALRAEFV	CCGGAGGTCCTGCGCGATGGTCAGGGCGAAGCTGCGCGCCGCCGGGCAGGCCGAGGCG	PEVLRGMVRAKLRAAGQAEA	GCAGGGCCGAACGTGGTCGACCGGCTCGCTCGGTCCGAGTCCGACCAGGTCGCGGGGGGGG	AGPNVVDRLAGRSESDQVAG	CTGGCCGAACTGGTGCGTTCACACGCGGCGGCGGTCTCCGGGTACGGCTCGGCCGACCAG
	22603		22663		22723		22783		2284		2290

					61/	89		(	)   		
	23022	ş	23082		23142		23202	i I	F16.2-		
LAELVRSHAAAVSGYGSADQ	CTCCCCGAGCGCAAGGCGTTCAAGGACCTCGGTTTCGACTCGCTGGCCGCGCGGTGGAGCTG	LPERKDLGFDSLAAVEL	CGCAACCGCCTCGGTACCGCGACCGGCGTGCGCTGCCCAGCACGTTGGTGTTCGACCAC	RNRLGTATGVRLPSTLVFDH	CCGACTCCGCTGGCGGTGGCCGAACACCTGCGGGACAGGCTGTTCGCGGCCTCACCGGCG	PTPLAVAEHLRDRLFAASPA	GTGGACATCGGCGACCGGCTGGAGAGGCGCCTCGAAGCCCTGTCCGCGAG	V D I G D R L D E L E K A L E A L S A E	GACGGGCACGACGTGGGCCAGCGCCTGGAGTCGCTGCTGCGCCGGTGGAACAGCAGG	DGHDDVGQRLESLLRRWNSR	CÉGGGGGGACGCCCGACGTCCGCGATCAGCGAGGACGCCAGTGACGACGAGCTGTTC
	22963		23023		23083		23143		23203		

					62	2/89			() 1	— О I	٠
23322		23382	i	23442	į	23502	ţ	23562	C	716.Z 23622	t
+	RADAPSTSAISEDASDDELF	TCGATGCTCGACCAGCGGTTCGGCGGGGGGGGGACCTGTAGATGAGCGGTGACACGGC	SMLDQRFGGEDL*MSGDNG	ATGACCGAGGAAAAGCTCCGGCGCTACCTCAAGCGCACCGTCACCGAGCTCGACTCGGTG	M T E E K L R R T L T E L D S V	ACCGCGCGCCTGCGTGAAGTCGAGCACCGGGCCGGTGAGCCGATCGCGATCGTCGGCATG	TAREVEHRAGEPIAIVGM	GCGTGCCGGTTCCCCGGCGACGTGGACTCGCCGGAGTTCGTTC	ACREPGDVDSPESFWEFVSG	GGCGGGGACGCCATCGCGGCCCCCCCCGCGGCGGCTGGGAGCCGGACCCCGACGCG	G G D A I A E A D A G W E P D A
23263		23323		23383		23443		23503		23563	

						.2-62					
23682	ı	23742	ı	23802	l	23862	1	23922	i I	G F16. - 23982	i
CGGCTGGGCGGGTGCTCGCGGCCGCGGCGACTTCGACGCGGGCTTCTTCGGGATCTCG	R L G G M L A A G D F D A G F F G I S	CCGCGCGAGGCGCTGGCGATGGACCCGCAGCAGCGGGATCATGCTGGAGATCTCGTGGAGAGAGA	PREALAMDPQQRIMLEISWE	GCGCTGGAGCGCCGGCCACGATCCGGTGTCCCTGCGCGGCAGCGCGACCGGGGTGTTC	A LERAGHDPVSLRGSATGVF	ACCGGTGTCGGCACCGTGGACTACGGCCCGCGACCGACGACGCCCCGGACGACGACCTG	T G V G T V D Y G P R P D E A P D E V L	GGCTACGTCGGCACCGCCTCCAGCGTCGCCTCCGGCCGGGTCGCCTACTGCCTG	G Y V G T A S S V A S G R V A Y C L	GGCCTGGAAGGCCCGGCGATGACCGTCGACACCGCCTGTTCCTCCGGGCTCACCGCCCT	GLEGPAMTVDTACSSGLTAL
23623		23683		23743		23803		23863		2392	) `

					-63					
24042	1	24102	I	24162	ì	24222	ł	24282		716.Z 24342
CACCTGGCGATGGAGTCGCTGCGCCGGGACGAGTGCGGCCTGGCCCTGGCCGGCGGCGTG	HLAMESLRRDECGLALAGGV	ACGGTGATGAGCAGTCCCGGGCGTTCACCGAGTTCCGCAGCCAGGGCGGGC	TVMSSPGAFTEFRSQGGLAA	GACGGCCGCTGCAAGCCGTTCTCGAAGGCCGCCGACGGGTTCGGCCTGGCCGAGGGTGCC	DGRCKPFSKAADGFGLAEGA	GGGGTCCTGGTGCTGCAACGGCTGTCGGCCGCGCGGGGGGGG	GVLVLQRLSAARREGRPVLA	GTGCTGCGGGCTCGGCGACCAGGCGCCCCAGCAACGGGCTGACCGCCCGAGC 3+++++++	VLRGSAVNQDGASNGLTAPS	GGACCCGCGCAGCAGCGGGTCATCCGCCGGGCGCTGGAGAACGCCGGTGTCCGGGCGGG
23983		24043		24103		24163		24223		24283

					6	5/89				Č	4
I	24402	l	24462	1	24522	1	24582	i	24642	(	G F 16.2- - 24702
GPAQQRVIRALENAGVRAG	GACGTCGACTACGTGGAGGCCCACGGCACCGGCACCAGGCTGGGCGACCCCATCGAGGTG	DVDYVEAHGTGTRLGDPIEV	CACGCGCTGCTCTCGACCTACGCGCGCGCGCGACCCCGGACGATCCACTGTGGATCGGT	HALLSTYGAERDPÓDPLWIG	TCGGTCAAGTCCAACATTGGCCACACCCCAGGCCGCCGCCGGCGTCGCCGGGGGTGAAG	SVKSNIGHTQAAAGVAK	GCGGTGCTGGCGCTGCGGCGAGATGCCGCGCACGCTGCACTTCGACGAGCCTCG	. >	CCGCAGATCGAGTGGACCTGGGCGCGGTGTCGGTGTCGCAGGCGCGCGGTGTGGCC	PQIEWDLGAVSVVSQARP	GCCGGCGAGAGGCCCCCCAGGGCGGCGTCTCCTCGTTCGGCATCAGCGGCACCAACGC
	24343		24403		24463			C 4 C F 7	24583	•	24643

					66	/89					0
1	24762		24822		24882	t	24942	1	25002		716.2 25062
GERPRAGVSSFGISGTNA	CACGTCATCGTCGAAGAGGCCCCGAGGCCGACGAGCCCGAGCCGGCACCCGACTCGGT	VIVEERAPEPEPAPDSG	CCGGTCCCGCTGGTGTTGTCCGGCCGCGACGAGCAGGCGATGCGGGCGCGCAGGCGGGACGG	VPLVLSGRDEQAMRAQAGR	CTGGCAGACCACCTCGCCGCGAGCCGCGGAACTCGTTGCGCGACACCGGTtTCACGCTG	ADHLAREPRNSLRDTGFTL	GCCACCCGCCGCAGCGCACCGCGCGCGGTGGTCGGCGACCGCGACGACGCCCCCCCC	TRRSAWEHRAVVVGDRDDA	CTCGCCGGGCTGCGCGGCCGACGGCCGCATCGCCGACCGGACGGCCACCGGGCAG	AGLRAVADGRIADRTAGQ	GCCCGAACTCGCCGCGCGTCGCGTGTTCCCCCGGCCAGGGCGCGCAGTGGCAGGGGG
4	CA 24703	I	. cc 24763	<b>Δ</b>	CT 24823	I	GC 24883	A	Z4943	H	G(25003 -

衣

					67/89			נ	Ω Ω 1	
1	25122	ì	25182	- t	25242	t	25302		FIG. 2-255362	•
ARTRGVAMVEPGQGAQGG	ATGGCCCGCGACCTGCGGGAGTCGCAGGTATTCGCCGACTCGATCCGCGACTGCGAG	MARDLLRESQVFADSIRDCE	CGGGCGCTGCCCCCCCACGTCGACTGGTCGCTGACCGACC	RALAPHVDWSLTDLLSGAP	CTGGACCGGGTCGTCGTCCAGCCCGCGCTCTTCGCCGTCATGGTCGCTGCGGCG	LDRVDVVQPALFAVMVSLAA	CIGIGGCGCICCCACGGCGICGAGCCCGCCGCGGTCGICGGCCACICGCAGGGGCGAGAIC	LWRSHGVEPAAVVGHSQGEI	GCCGCCGCGCACGTCGCCGCGCGCTCACCCTGGAGGACGCCGCCAAGCTCGTCGCGGTC	A A H V A G A L T L E D A A K L V A V
	25063		25123		25183		25243		25303	

CGTTCGGGC.	GGGCGCGCTCCATCGCTCGGTC++++	GGGCCGCTGGACGAGCTGATCGCC -+++++	CCCCGTCGACTACGCCTCCACTCA+++	CCGAGCTGGCGTCTCCCCGGTG++++++++	CAGCCCATCGACCGCCACGATG -++++++
SGCTCGGCGGCCAGGC	ACCGAGCGGCCGAACGGATCGGGCGCTTCGCGG(25423++++	AACGGCCCCCGGTCGTCGTCGCGGGGAGAGCGGCCGCT 25483++++++++	GAGTGCGAGGCCGAAGGCATAACGGCGCGCGCGCATCC	CCGCAGGTGGAGTCGCTGCGCGAGGAGCTGCTGACC	TCGGCGACGTGGCGCTCTACTCGACCACGACCGGGG

					69	/89					Ω Ω 
1	25782	ì	25842		25902	I	25962	1	26022	( i	F1G.Z 26082
SADVALYSTTGQPIDTATM	GACACCGCCTACTGCGCAACCTGCGCGAGCAGGTCCGCTTCCAGGACGCGACGCGG	DTAYWYANLREQVRFGDATR	CAGCTCGCCGAGGCGGGTTCGACGCGTTCGTCGAGGTCAGCCCGCATCCGGTGCTGACC	QLAEAGFDAFVEVSPHPVLT	GTCGGCATCGAGGCCACGCTGGACTCCGCGCTCCCGGCCGACGCCGGCGCCTGCGTCGTG	VGIEATLDSALPADAGAVV	GGCACCCTGCGGGGACCGCGGCGGCCTGGCCGACTTCCACACCGCGCTCGGCGAGGCG	GTLADFHTALGEA	TACGCGCAGGGCGTGGACTGGAGCCCCCCCCTTCGCCGACGCGCGGCCGGTCGAG	YAQGVEVDWSPAFADARPVE	CIGCCGICTACCCGITCCAGCGGCAGCGGTACTGGCTGCCCATCCCCACCGGCGGGCG
	25723		25783		25843		25903		25963		26023

					70	/89				Ċ	D 0 1
	26142		26202	1	26262	i	26322	1	26382		F16.2 26442
LPVYPFQRYWLPIPTGGR	GCACGGGACGACGACTGGCGCTACCAGGTCGTATGGCGGGAAGCCGAGTGGGAG	ARDEDDWRYQVVWREARE	AGCGCTTCGCTGCCGGACGCGTGCTGCTGACCGGACCGG	SASLAGRULLVTGPGVPSL	TCGGACGCCATCCGAAGTGGACTGGAGCAGAGCGGTGCGACGGTCCTGACCTGCGACGTG 3+++++++	SDAIRSGLEQSGATVLTCDV	GAATCCCGTTCGACCATCGGCACCGCACTGGAGGCCGCCGACCCGACGCTCTGTCCACT 3++++++++	ESRSTIGTALEAADTDALST	GIGGIGICGCIGCTGTCCCGCGACGGCGACGCCGTCGATCCGTCGCTGGACGCGTCGCC 3+++++++	V V S L L S R D G E A V D P S L D A L A	CTGGTCCAGGCCCTCGGAGCGGCCGGGGTCGAAGCACCGCTGTGGGTGCTGACCCGCAAC
	26083		26143		26203		26263		2632		26383

	·				71/	89		1	O / I		
1	TC 26502	1	3AC + 26562	1	GAG + 26622	l Ei	cgc + 26682		FIGGG +16.Z	l G	SCTG
V Q A L G A A G V E A P L W V L T R N	GCCGTGCAGGTGGCCGACGGCGAACTGGTCGATCCGGCGCGCAGGCCATGGTGGGCGGTC	V Q V A D G E L V D P A Q A M V G G L	GGCCGCGTGGTCGGCATCGAGCCGGGGCGCTGGGGCGGCGGTCTGGTGGACCTGGTCGAC	R V V G I E Q P G R W G G L V D' L V E	GCCGATGCCGCGTCGATCCGGTCGCTGCCGCGGTGCTGGCGGGGGGGCGCGCGC	DAASIRSLAAVLADPRGE	CAGGTCGCGATCGGGGGGGGGTGGCGAGGCTCGTGCCCGCCC	V A I R A D G I K V A R L V P A P A	GCCGCACGCACCCGCTGGAGCCCTCGCGGCACCGTGCTGGTCACCGGCGGCACCGGAGGG	ARTRWSPRGTVLVTGTG	ATCGGCGCGCACGTCGCCTGGCTGGCCCGCTCGGGCGCCGAGCACCTGGTGCTGCT
IJ	GC( 26443	K	GG( 26503	U	GC 26563	A	CA 26623	Ø	GC 26683	A	A

						72/89			ì		
									(	-7.	
26802	i	26862	I	26922	ı	26982	ı	27042	ı	F1G.	1
+	IGAHVARRIASGAEHLVLL	GGCAGGCGCGCCGACGCACCCGGCGCGTCCGAGCTGAGGGAGG	GRRGADAPGASELREELTAL	GGCACGGGCGTGACCATCGCCGCCTGCGACGTCGCCGACCGGGCGCGCGGCTCGAAGCGGTG	GTGVTIAACDVADRARLEAV	CTCGCCGCGGAGCGCACGCACGGTCAGCGCCGTGATGCACGCGGGGGGTTT	LAAERAEGRIVSAVMHAAGV	TCCACGTCCACGCCCCTCACCTCACCGAAGCCGAGTTCACCGAGATCGCCGACGTG	STSTPLDDLTEAEFTEIADV	AAGGTGCGCGCACCGTCAACCTGGACGAGCTCTGCCCGGGACCTCGACGCGTTCGTT	K V R G T V N L D E L C P D L D A F V L
26743		26803		26863		26923		26983	) 	27043	

ຮ

•		7	73/89		7
27162	27222	27282	27342	27402	F16.2.
TTCTCCTCCAACGCGGCGGCAGTCCGGGGCTCGCCTCCTACGCGGCGGCCAAC+++++++++-	TCCTC	GCCTGGGGCTCTGGGCCGGGCAACATGGCCGGGGACGAGGCGGCGAGTACCTGCGC+++++++-	AGCCAGGGCCTGCGGGCCATGGACCCGGATCGGGCCGTCGAGGAACTGCACATCACCCTC++++++	GACCACGGTCAGACGTCGTGGACATGGATCGCAGGCGGTTCGTCGAGCT(++++++	TTCACCGCGCCCGGCCGCTGTTCGACGAGATCGCCGGTGCCCGGGCGGAAGC+++++-
27103	27163	27223	27283	27343	27403

Ė

-					7	4/89			1	5/-7	-
GTCGACGGCCGAG -+	STAE	rgcrcgccacgc -+27582	L G H G L	TCGACTCCATGACC -+27642	D S M T	AAGCCGCGACGGTG -+27702	AATV	AGCGGCTCGTCGGC +	R L V G	AGGATGCCGACGAC FIG.Z	- Q Q Y
CTCGCGGCGC7	L A A L	GTCGCCGCGG	V A A V	CGACCTCGGCT	D L G F	CGGGGTGCGGG	G V R	CACTACCTGG	H Y L	CAGGTGCCGA	E V P
CAGCGGC	<b>K</b>	CCGAG	回	CITCCGC	E K	GGTGACC	H >	GCCGAC	. A	GTGCGC	۷ ه
-+	Ø	CCGCG(	R	CGC + -	A A	+	æ	GGCTC	ы	CGCTC	ы
SCGCT	. L	CTGAT	н	CGACCG	D R	GCTCG(	LA	ACCC	F K	20000	ъ Д
+	<u>م</u>		I	ACCGC-+-	r R	ACCGG -+	ፈ	CCATC	н	AAGCC	A C
GAGGG	២	CTCGC	I. A	BATCG	П	SCGGA	Z X	CCCGA	Д E+	GGAGC	EQ C
CGAG	ជ	AGCAC +	Ħ	2000c	K	ACCTC +	H	SACCA(	Ħ	ĠAGGC -+	A.
CGGCAGAGCGAGGGGCCCGG	S	AGGCGCGAGCACCTCGCCCAC	다 더	GACGACGCGCGATCGACCG	D	GCCGTCGACCTGCGGAACCG	Δ >	GTCTTCGACCACCCGACCAT	ra Ö	CAGCA	A E
CG 27463	K	AG 27523	K.	G2 27583	Q	G(27643 -	K	G 27703 -		G 27763 -	Æ

Ē

15

•						7	<b>1</b>		
27882	27942	1	28002	i	28062	i	28122	ı	rib. 28182
CCGATCGCGATCGTCGGCATGGCCTGCCGCTTCCCCGGCGGCGTGCACAACCCCGGT +++++	PIAIL AIVGMACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	I W	TGGGACCTCGACGCGCTGTTCGACCCCGACCCGCACGCCCACGGAACCAGCTACTCGCGA	WDLDALFDPPQRHGTSYSR	CACGGCGCGTTCCTCGACGGGCCGCCGACTTCGACGCGGCGTTCTTCGGGATCTCGCCG	HGAFLDGAPPFDAAFFGISP	CGCGAGGCGCTGGCGATGGACCCGCAGCAGCCCAGGTCCTGGAAACGACGTGGGAGCTG	REALAMDPQQRQVLETTWEL	TTCGAGAACGCCGGCATCGCCCCCCCCTCGCTGCGGGCAGCGACACCGGCGTCTTCCTC
27823	27883		27943		28003		28063		28123

					76	/89				7	0/1
1	28242	ŧ	28302	1	28362	ı	28422	i	28482		F16.2 28542
EENAGIDPHSLRGSDTGVFL	GGCGCCGCGTACCAGGCTACGGCCAGGACGCGGTGGTGCCCGAGGACAGCGAGGGCTAC	GAAYQGYGDAVVPEDSEGY	CTGCTCACCGGCAACTCCTCCGCCGTGGTGTCCGGCCGGGTCGCCTACGTGCTGGGGCTG	LLTGNSSAVVSGRVAYVLGL	GAAGGCCCCGCGGTCACGGTGCACGGCGTGTTCGTCGTCGTTGGTGGCCTTGCATTCG +++++++	EGPAVTVDTACSSSLVALHS	GCGTGTGGGTCGTTGCGTGACTGCGGTCTTGCGGTGGCCGGTGGTGTGTTCGGTG	A C G S L R D G D C G L A V A G G V S V	ATGGCGGCCCCGGAGGTGTTCACCGAGTTCTCCCGCCAGGGCGGGC	MAGPEVFTEFSROGGLAVDG	CGCTGCAAGGCGTTCTCCGCGGAGGCCGACGGCTTCGGTTTCGCCGAGGGCGTCGCGGTG
	28183		28243		28303		28363		28423	•	28483

				77/89			1	9/-7		
I	28602	1	28662	l	28722	1	28782	i	F16.2	t
R C K A F S A E A D G F G F A E G V A V	GTCCTGCTCCAGCGGTTGTCCGACGCCCGCAGGGCGGGTCGCCAGGTGCTCGGCGTGTC	VLLQRLSDARRAGRQVLGVV	GCGGGCTCGGCGATCAACCAGGACGCGCGAGCAACGGTCTCGCGGCGCCGAGCGGCGTC	A G S A I N Q D G A S N G L A A P S G V	GCCCAGCAGCGCGTGATCCGCAAGGCGTGGGCGCGTGCGGGGGATCACCGGCGCGCGATGTG	A Q Q R V I R K A W A R A G I T G A D V	GCCGTGGTGGAGGCGCATGGGACCGGTACGCGGCTGGGCGATCCGGTGGAGGCGTCGGCG	AVVEAHGTGTRLGDPVEASA	TIGCIGGCIACITACGGCAAGICGCGGGGTCGICGGGCCCGGTGCTGCTGGTTCGGTG	L A T Y G
<b>♥</b> - 10 <b>♥</b>	28543		28603		28663		28723	·	Q	C0 / 0 7

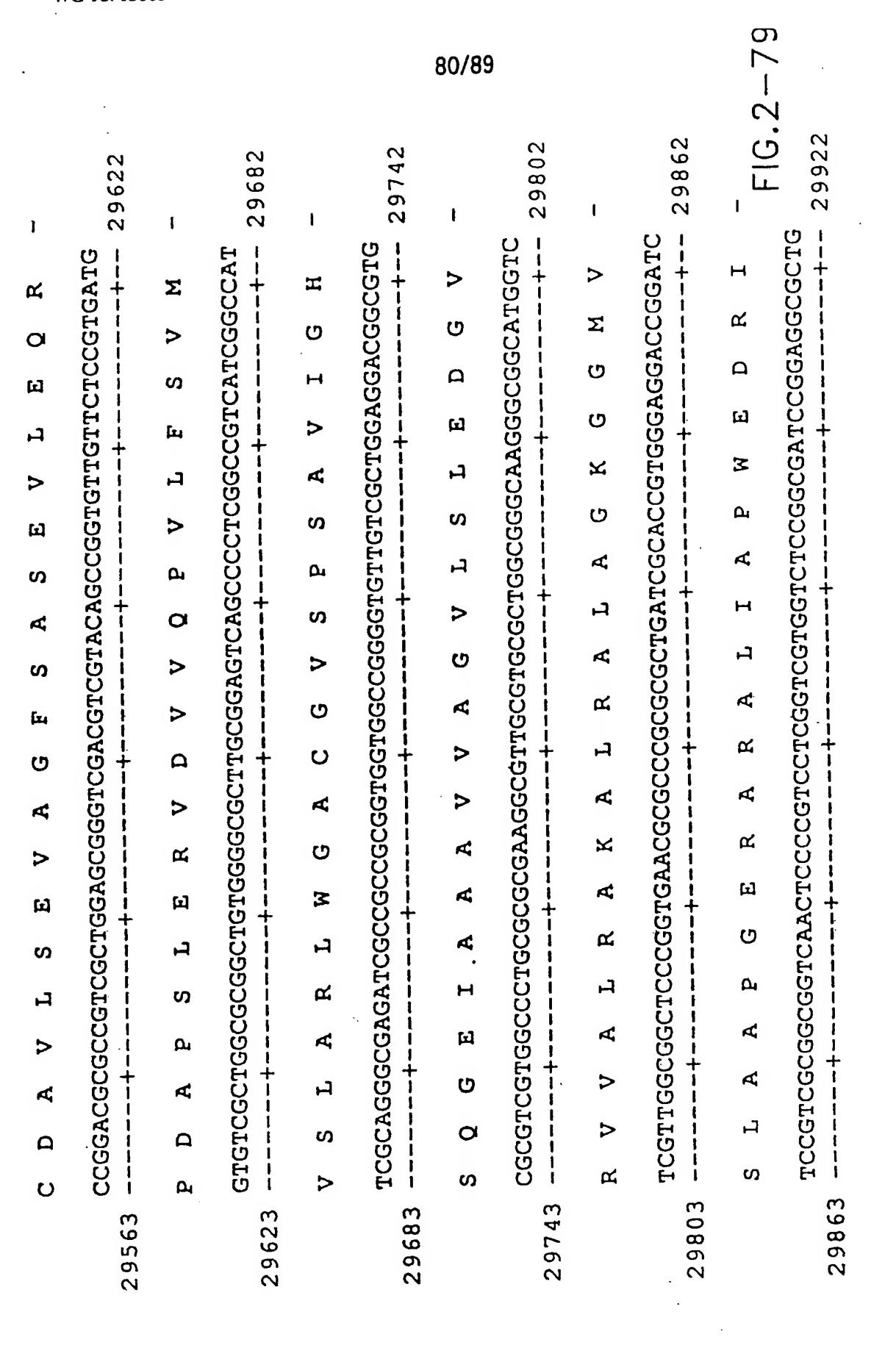
Ě

·	•	78/8	9		/ / - 7
28902	28962	29022	29082	29142	F16.
AAGTCGAACATCGGTCACGCGGCGGCGGCGGGCGTGATCAAGGTGTC+++++++	CTGGGGTTGAACCGCGGCCTGGTGCTCTGCCGCGGCGAGCGGTCGCCGTG+++++++	ATCGAATGGTCCTCGGGTGTGTGGAACTTGCCGAGGCCGTGAGCCCGTGGCCTCCGGCC++++++	GCGGACGGGCCGGGCCGGTGTCGGCGTTCGGGGTGAGCGGGACGAACGCGCAC  3+++++++	GTGATCATCGCCGAGCCCGAGCCGCTGCCGGAACCCGGACCGGTGGGCGTG 3+++++++	CTGGCCGCTGCGAACTCGGTGCCCGTACTGCTGTCGGCCAGGACCGAGACCGCGTTGGCA
28843	28903	28963	29023	29083	. 29143

				7		1	Ω 				
•	29262	1	29322	1	29382	1	29442	1	29502	i	F16.2
A A A N S V P V L L S A R T E T A L A	CGCAGGCGCGCTCCTGGAGTCCCGCAGTGGACGACTCGGTTCCGTTGACGGCATTGGCT	A Q A R L L E S A V D D S V P L T A L A	TCCGCGCTGGCCACCGCGCCCTGCCGCGTCGTGCGGCGTTGCTGGCAGGCGAC	SALATGRAHLPRRAALLAGD	CACGAACAGCTCCGCGGCAGTTGCGAGCGGTCGCCGAGGGCGTTGCGGCTCCCGGTGCC	HEQLRRAVAEGVAAPGA	ACCACCGGAACCGCCTCCGCCGGCGTGGTTTTCGTCTTCCCAGGTCAGGGTGTTCAG 	T'T GTASAGGVVFFVFPGQGAQ	TGGGAGGGCATGGCCCGGGGCTTCTCGGTCCCCGTCTTCGCCGAGTCGATCGCCGAG	WEGMARGLLS VPVFAESIAE	TGCGATGCGGTGTTCTGGCCGGGTTCTCGGCCTCCGAAGTGCTGGAGCAGCGT
H	G 29203 -	~	29263		29323		29383		29443	•	29503

۲.

3



Ď

**?** ·

					•	81 /	<b>′8</b> 9	(	08-		
1	29982	i	30042	1	30102	1	30162	( (	FIG.2-	ľ	
S V A A V N S P E A L	GCCGAACTCGTCGCACGAGGACGAGGGCGTGCGCGCCCAAGACGCTCCCGGTGGAC	AELVARCEDEGVRAKTLPVD	TACGCCTCGCACTCCGCCACGTCGAGGAGATCCGCGAGACGATCCTCGCCGACCTCGAC 3+++++++	YASHRUEEIRETILADLD	GGCATCTCCGCGCGCGTGCCGCCATCCCGCTCTACTCCACGCTGCACGGCGAACGGCGCG	GISARAAIPLYSTLHGERRR	GACGGCGCCGACATGGGTCCGCGTACTGGTACGACAACCTGCGCTCCCAGGTGCGCTTC	DGADMGPRYWYDNLRSQVRF	GACGAGGCGGTCTCGGCCGCCGTCGCCGTCACGCCACCTTCGTCGAGATGAGCCCG	DEAVSAPDGHATFVEMSP	CACCGGTGCTCACCGGGGGGGGGAGATCGCCGCGGGGCGGTGGCCATCGGGTCG
*	29923		29983		30043		30103		30163		

SUBSTITUTE SHEET

						82/	89		С	D I	
										V	
82		42		02		162		522		<u>↑</u>	
3028	1	30342	·	30402	•	304	i	308	ı	308	I
1.,	•		•		•			Ö I		Ω I	
+	w ·	3CA -+-	Ħ	CAA -+-	Z	CGAC -+	Ω	000	Ö	AG( -+-	Ů
1	O	GT(	>	Ü	Ωı	000	K	16G	G	990	O
1	Н	CAC	I	CHG	ы	CTC	H	GA	ы	960	A
	Æ	3C G	Æ	9 I	Æ	CAG	a	CTG	그	AAG	×
+	>	CCGGCGCACGTGCAC	æ	GGTGGCGCTGCCCAAC	>	CGACCAGCTCGC	Ω	GGACCTGGAAGGCGG	Ω	rcgagaaggccggagg	ជ
	4	ŬΪ		O 1	Д	ICC I	တ	<b>⊢</b>	>	(b)	>
! !	K	TCG	K	CTC		H.C.		2000	•	GCA	
į	Ω	Je C	H	ZAC FF-	Ωι	AGG +	>	) - + + - +	Ωı	ŬΙ	Æ
+	K	CGAG	臼	+ + 000	K	+	问	CAC	H	CAG +	S
1	Æ	292	A	099	K	00	Ω	CAC	H	GAC	H
1	н	GAT(	н	CCC	<u>ρ</u>	090	A.	099	K	GCT	н
+	阳	CT(-+-	H	CHH	<b>चि</b>	ĞCT -+-	니	- + -	H	GTC	ഗ
	a	3CA(	耳	CGT(	>	CTG	3	AC C	Δ,	GGA	ធ
	>	3GA(	臼	GAA(	Z	GTA	>-	909	民	ACC	Ωı
1	æ	GGA( +	臼	GCG(++	K	+-	R.	CTG +	Z	+ +	K
	<b>4</b> .	)    - 	A	(U)	Z		ø	CGA	Ω	GTC	လ
	ĘH		H	GGCGTGGCCGTGGACT	Ω	TACCCGTTCGAGCCCC	а	AGCCGCTACCGCGTCG	>	TTCCTGGTCCACGGGT	Ŋ
 	H	CTGCACCGCGACAC	Ω	CGT	>	CGA	臼	500	民	CCA	Ħ
<del> </del>  -	>	) CG +	<b>1</b>	3.GC	A.	H H H	ĮΞι	CTA +	>1	TOUT	>
t t	Ďι	CAC	#	GTO	>	Ö	<b>ב</b>	Ö	<b>K</b>	D I	ы
	Ξ.	CTG	ı	299	ŋ	TAC	≯	AGC	S	TTC	Įτι
30223 -	•••	30283		30343		30403		30463		30523	

12		02		62		83/89		82	) O O O O O	-10.6 942	
30642	i	30702	ı	307	1	308	ŧ	30882	ı	308	ī
CGCGTCGTGCCGGCCTCGCCGACCGCGAAGCCTCGGCGGCCCTGCGGGAGGTGCCG	R V V P V A S A D R E A S A A L R E V P	GGCGAGGTCGCCGGCGTCTCGGTCCACCCGGCGCCGCAACGCACCTCGCCTGCAC	GEVAGVLSVHTGAATHLALH	CAGTCGCTGGGTGAGGCCGGCGGGCCCCCCCTCTGGCTGG	QSLGEAGVRAPLWLVTSRAV	GCGCTCGGGGAGTCGATCCCGAGCAGCGGCGATGGTGTGGGGTCTCGGGCGCGC	A L G E S E P V D P E Q A M V W G L G R	GTCATGGGCCTGGAGCCCCGGAACGGTGGGCGGTCTGGTGGACCTGCCCGCCGAACCC	V M G L E T P E R W G G L V D L P A E P	GCGCCGGGGGACGCGAGGCGTTCGTCGCCTCCCTCGCGCGCG	APGDGEAFVACLGADGHEDQ
30583		30643		30703		30763		30823		30883	

Ą

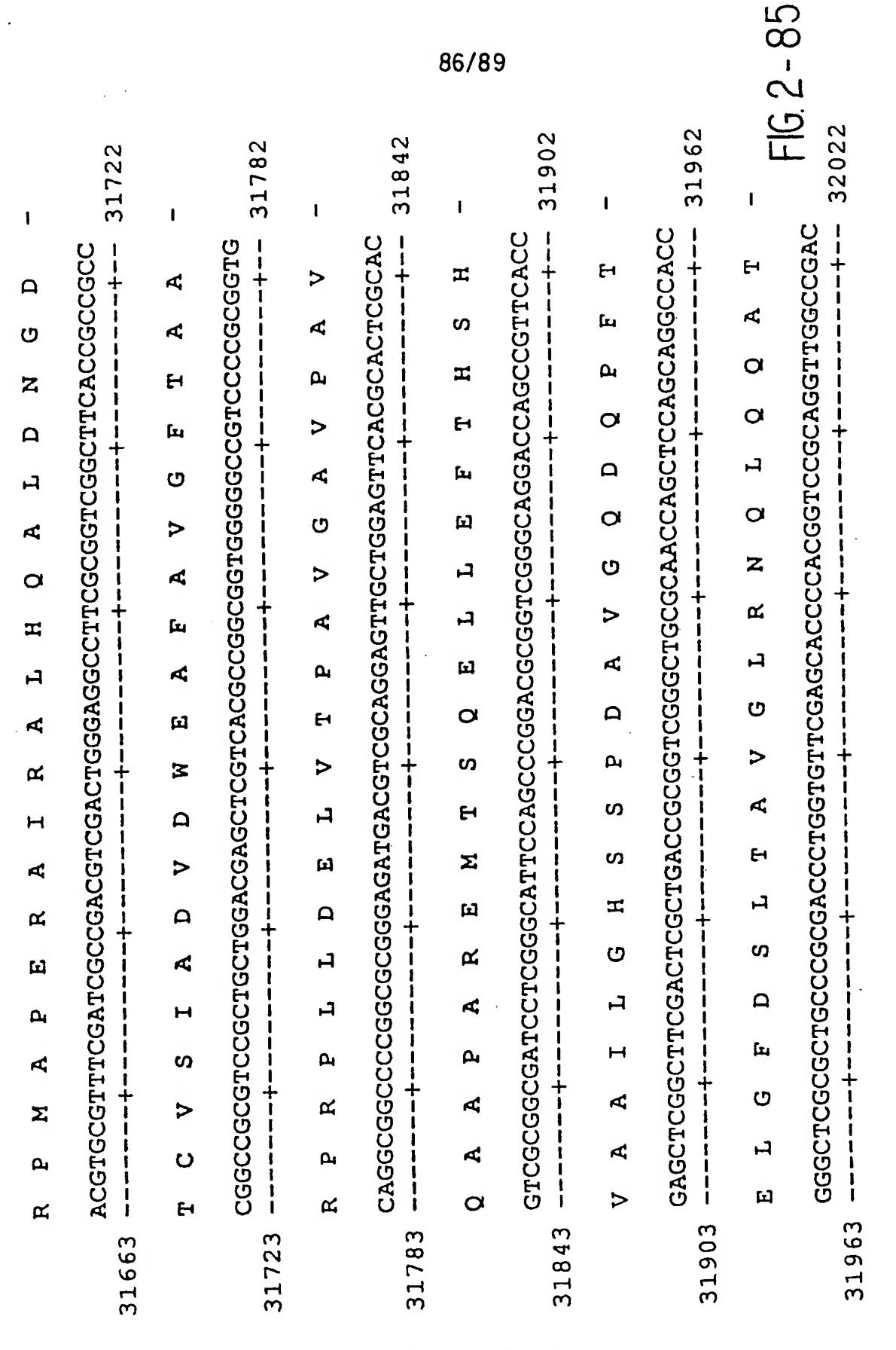
		·		84/89				2-83		
31002	I	31062	l	31122	1	31182	i	31242	6 51d _	33
GTCGCGATCCGTGACCACGCCCGCTACGGCCGCCGCCTCGTCCGCGCCCCCGCTGGGCACC	VAIRDHARYGRRLVRAPLGT	CGCGAGTCGAGCTGGGAGCCGGGCACGGCGCTGGTCACCGGCGGCACCGGTGCGCTC	RESSWEPAGTALVTGGTGAL	GGCGGCCACGTCGCCACCTCGCCAGGTGCGGGGTGGAGGACCTGGTGCTGGTCAGC	G G H V A R H L A R C G V E D L V L V S	AGGCGCGCGTCGACGCTCCCGGCGCCGAGCTGGAAGCCGAACTGGTCGCCTCGGC	RRGVDAPGAAELEAELVALG	GCGAAGACGACCATCACCGCCTGCGACGTGGCCGACCGCGAGCAGCTCTCCAAGCTGCTC	A K T T I T A C D V A D R E Q L S K L L	GAAGAACTGCGGGGCAGGGACGTCCGGTGCGGACCGTCGTGCACACCGCCGGGGTGCCC
30943		31003		31063		31123		31183		31243

7

					{	35 <b>/</b> 8:	9			70	† 0
	01		<b>~</b>		7		8		2	00	. 7.0.C 62
ı	31362	ı	31422	ŧ	3148	1	3154	1	31602	1	316
ELRGQGRPVRTVVHTAGVP-	GAATCGAGGCCGCTGCACGAGATCGGCGAGCTGGAGTCGGTCTGCGCGCGC	SRPLHEIGELESVCAAKVT	GGGGCCCGGCTGCTCGACGTGTGCCCGGACGCCGAGACCTTCGTCCTGTTCTCGTCC+++++++	ARLLDELCPDAETFVLFSS-	GGAGCGGGGGTGTGGGCAACCTCGGCGCCTACTCCGCGGCCAACGCCTACTC+++++++	A G V W G S AN L G A Y S A AN A Y L	GACGCGCTGGCCCACCGCCGTGCGGAAGGCCGTGCGGCGACGTCCGTC	ALAHRRAEGRAATSVAWG-	GCCTGGGCGGCGAGGGCATGGCCACCGGCGACCTCGAGGGGCTCACCCGGCGCGCGC	WAGEGMATGDLEGLTRRGL.	CGCCCGATGGCGCCCGAGCGCGCGCGCGCGCTGCACCCAGGCGCTGGACAACGGCGACACGCGACACGGCCGACACGGCCGACACGGCCGACACGGCCGAC
田	GA 31303	时	GG 31363	<b>o</b>	31423		GA 31483	Q	31543	K	31603

**'t**,

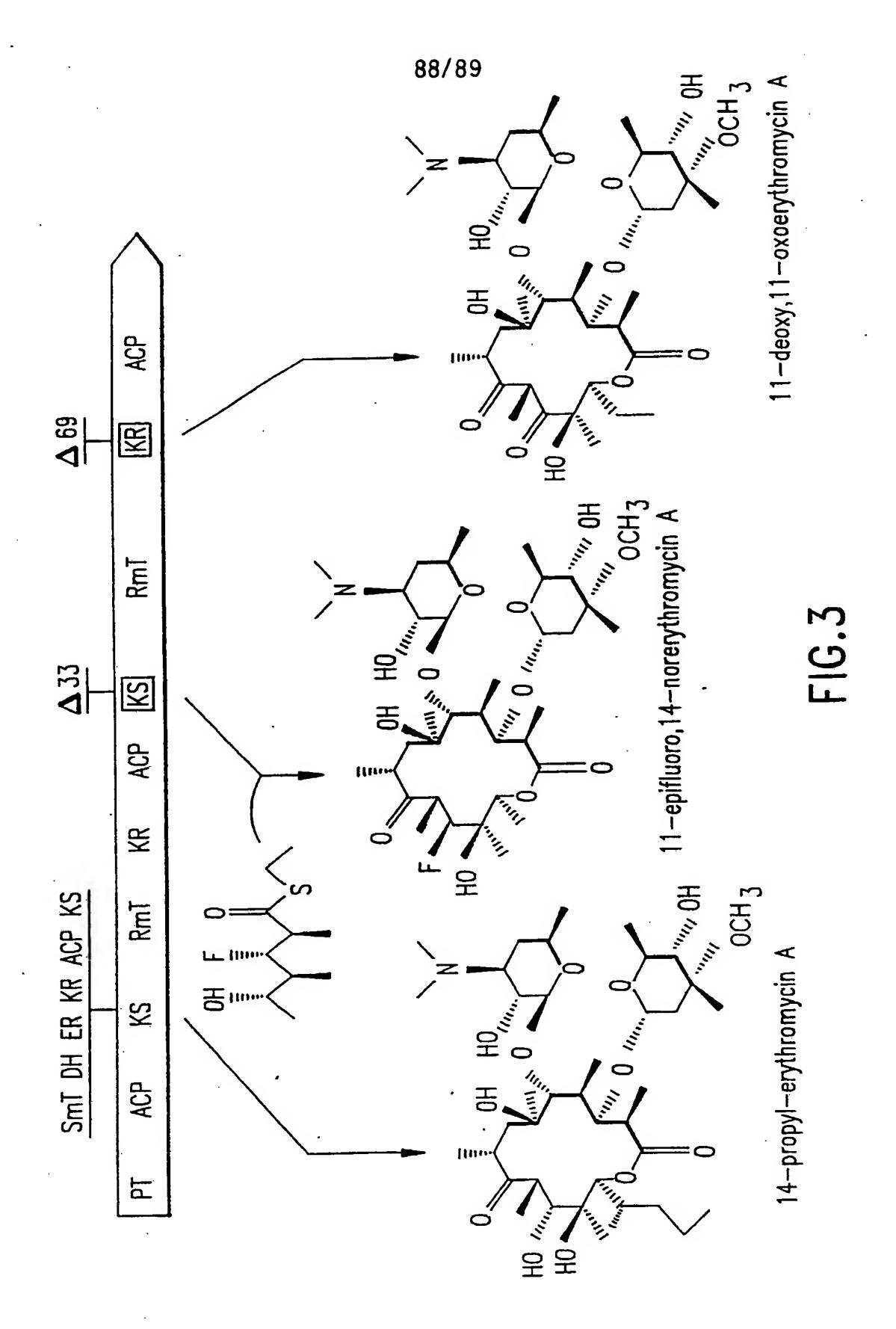
**3**.



			87/89						$\bar{\alpha}$	
									F16.2	
	82		42		02	:	62		<u></u>	
	32082	1	321	ı	32202	i	322	1	F1	i
•					۵ ۱		H L		99	<b>.</b> 4
Ω	TCC +	R.	) 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	K	rcgrg +	>	CGA +	H	H H H	CC.
K	TCI	H,	, GC	H	CHA	.H.	) 1 1 1	K	000 1100	<b>&gt;</b>
Н	AGCAGCGCTCTTCGC ++	K	CTCGACCTGCTGGCG	н	TCCCTCGATCT	Ω	GGAACGGCGGCGATC ++	A	CTCCGGTTCGG	Ω
K	CAG +	တ	CGA +	Ω	+ CT	ы	AAC +	EH	ATCGC	K
民	GAG	ഗ		н	CHO	ഗ	999	ပ	AAT	Н
>		æ	CTA(	×	GTT	[Li		A	090	Ŋ
H	CCGGGAA	ធ	O	ဟ	CGACGG	U	CTG	O	525	ፈ
Δı	) +	R.	CCGGT	K	CGA +	Ω	GATCTGCT	U	GCT(	H
耳	CGC I	æ	GGT	>	CGGCTC	ഗ	GAT	н	099	<b>A</b>
ы	O	Д	CAG	R.	990	U	GGT	>	992	ပ
[Ŧ1	CGGGACTC	EH	GTCGGGCAGGGT	ပ	CGA( -+-	Ω	CAC	E	CGC	K
>	2002	G	GTC	တ	HH	្រុ	GGT	>	GCT	н
Н	CAGO	တ	CGT(	>	GCA 	I	AGA(	ជ	500	K
EH	3GA(	Д	. – + + – – +	O	CGA(++	ធ	000 +	Ö	CAC +	E
Ø	SCTO	+	) 1 1 1 1	K	0 0 1	CC.	HCC I	ф	GTT	ŢĿı
ρι	SCAC	O	SCAC	Ø	HH	Ĺ	000	G	CGA	ជ
H	CACATAGGACAGCAGCTCGA	0 0 0	GACGGCTACCGGCAGGCGGG	<b>₽</b>	GGGCTGTCGGACTTCCG	S	GACATGGCCGACGGTCC	Ω	TCCGGTCCGCACGAGTT	I
E A	4GG2	Ü	CTAC	×	GTC( +:	ß	+	<b>A</b>	TCC0	Ωı
ы	ZAT2	н	100 100 110	ტ	SCT	ㅂ	CAT	Σ	990	Ů
Ü	CAO	I	GA(	Ω	999	<sub>ග</sub>	GA	Ω	Ö I	ທີ
	32023		32083		32143		32203		32263	

12

Þ



SUBSTITUTE SHEET

89/89

でと説

NUMBER  1 2 3 4 5 6 7 8 9	SITE BamHI Pvull Pvull Pstl BamHI Xhol Pstl Hindll Sphl EcoRl	DISTANCE (Kb) <sup>a</sup> -3.60  -3.50  -3.40  -3.05  -2.95  -2.80  -2.00  -1.60  -1.55  -1.50
7 8	Pstl Hindll Sphl EcoRl	-2.00 -1.60 -1.55 -1.50
11 12 13 14 15	Kpnl EcoRl Smal <sup>b</sup> Sphl Kpnl	-1.35 -1.05 -0.90 -0.75 -0.65
16	Smal	-0.20

FIG.4

## INTERNATIONAL SEARCH REPORT

International Application No. PCT/US92/00427

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)3									
	to International Patent Classification (IPC) or to bot								
IPC (5)	: A01N 43/22; A61K 31/71; C07H 17/08; : 435/76, 252.35, 886; 514/29; 536/7.	.2 C12N 1/21; C12P 19/62							
II. FIEL	OS SEARCHED	entation Secretary							
		entation Searched <sup>4</sup>	<del></del>						
Classificati	Off 2 Aproxis								
v.s.	435/76, 183, 252.35, 88	6; 514/29; 536/7.2							
	Documentation Searched to the extent that such Document	other than Minimum Documentation are included in the Fields Se	on erched <sup>6</sup>						
APS, BIOSIS, MEDLINE, BIOTECH. ABSTRACTS									
III. DOC	UMENTS CONSIDERED TO BE RELEVANT 14		·						
Category*	Citation of Document, 18 with indication, where app	ropriate, of the relevant passages 17	Relevant to Claim No. 18						
У	1 de la de la decimina del decimina de la decimina de la decimina del decimina de la decimina del decimina de la decimina del decimina de la decimina del del del decimina del								
Y	J. of Bacteriology, Volume 172, No. 5, issued May 1990, J.M. Weber et al, "Organization of a Cluster of Brythromycin Genes in Saccharomyces erythraea", pages 2372-2383, See the entire document.								
У	US, A, 4,874,748 (Katz et al) 17 October 1989, see 1-30 column 1, lines 41-59; column 3, lines 47-60.								
y	US, A, 4,935,340 (Baltz et al) 19 document.	June 1990, see entire	1-29						
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention document but published on or after the international filing date  "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  "O" document referring to an oral disclosure, use, exhibition or other means  "P" document published after the international filing date  "T" later document published after the international filing date and not in conflict with the application but cited to understand the principle or theory underlying the invention cannot be considered novel or cannot be considered to involve an inventive step document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art document member of the same patent family									
Date of the Actual Completion of the International Search 2  Date of Mailing of this International Search Report 2  30 MAR 1992									
	MARCH 1992	30 MAR 1992							
International Searching Authority  Signature of Authorized Officer 20  Dian Cook  Dian Cook									